PARAMETER PLANE ANALYSIS OF SAMPLED DATA SYSTEMS [WITH EXTENSIONS FOR CONTINUOUS SYSTEM APPLICABILITY]

ALOYSIUS R. MILLER

PARAMETER PLANE ANALYSIS OF SAMPLED DATA SYSTEMS

(WITH EXTENSIONS FOR CONTINUOUS SYSTEM APPLICABILITY)

* * * * *

Aloysius R. Miller

PARAMETER PLANE ANALYSIS OF SAMPLED DATA SYSTEMS

(WITH EXTENSIONS FOR CONTINUOUS SYSTEM APPLICABILITY)

by

Aloysius R. Miller

Lieutenant, United States Navy

Submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

IN

ELECTRICAL ENGINEERING

United States Naval Postgraduate School
Monterey, California

1 9 6 5

PARAMETER PLANE ANALYSIS OF SAMPLED DATA SYSTEMS

(WITH EXTENSIONS FOR CONTINUOUS SYSTEM APPLICABILITY)

by

Aloysius R. Miller

This work is accepted as fulfilling

the dissertation requirements for the degree of

DOCTOR OF PHILOSOPHY

IN

ELECTRICAL ENGINEERING

from the

United States Naval Postgraduate School

ABSTRACT

Many of the advances in servo-system design fostered by the general availability of digital computers are not utilized by practicing engineers because they discard rather than complement existing methods. An exception is the parameter plane method which retains the principle of design by such familiar specifications as damping and settling time, but where classical methods display the effect of only one variable on these specifications the parameter plane method can display the simultaneous effect of two. In this paper the restrictions on problems suitable for parameter plane analysis are eased and the number of specifications which can be considered on the parameter plane is enlarged to include bandwidth and steady state error. Parameter plane methods are also adapted to dominant mode design in a manner which not only allows selection of the dominant mode but simultaneously guarantees its dominance. These extensions of parameter plane theory are related to both continuous and sampled data systems.

ACKNOWLEDGEMENTS

The author expresses his appreciation to Dr. G. J. Thaler, Professor of Electrical Engineering at the United States Naval Postgraduate School for this topic and for the guidance provided in the preparation of this manuscript.

The author also thanks LT F. H. Hollister, USN for insights provided in his many helpful conversations.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE		PAGE

FIGURE PAGE

[illegible] [illegible]

[illegible] [illegible]

[illegible] [illegible] [illegible]

[illegible] [illegible] [illegible]

[illegible] [illegible] [illegible]

[illegible] [illegible] [illegible]

LIST OF TABLES

CHAPTER I

In recent years great emphasis has been placed on the "controls field" in electrical engineering. With this emphasis has come a proliferation of techniques for analysis and design of servomechanisms, many of which have not yet been fully exploited. Among such techniques are the parameter plane methods which can be extended in both applicability and in specifications considered.

Before analyzing the parameter plane as a design tool it may be wise to establish its place among tools presently available to the engineer. A broadbrush review of available methods should be sufficient for this purpose. Present day servomechanism design subdivides conveniently into two disciplines, direct time domain analysis and frequency or transform analysis. Of the two the former appears to be the more powerful albeit the more complex. Design by this procedure ordinarily includes the steps:

a) determination of signal characteristics
b) selection of a criterion by which system performance can be evaluated (i.e. integral squared error, etc.)

When the results of these steps are clear cut and definite time domain analysis will probably be the best tool available to the design engineer. This holds particularly for sampled data systems where incorporation of a digital computer can make realizability of theoretically obtained results a minor problem. Despite its power time domain analysis is by no means a panacea. By its very nature the time domain approach demands that system structure, its inputs and its objectives be well defined. These are by no means trivial requirements since a system designed on the basis of integral

squared error, for example, need not have acceptable acceleration characteristics, overshoot, frequency response, etc. Furthermore, analysis in the time domain generally provides little insight into possible performance criteria trade-offs or structural changes which might enhance the overall system.

The difficulties of obtaining performance characteristics from the integrodifferential equations which describe a servo-system have to a large extent been overcome by the use of transforms which reduce the operations of differentiation and integration to operations of multiplication and division. When transformed, the pole zero pattern or magnitude phase plot of the integrodifferential equation furnishes an indication of system performance by inspection. In addition these plots often indicate system modifications or compensation which will improve system performance. The primary disadvantage of transform analysis is that pole-zero patterns, bandwidth and other transform criteria abstract the time domain performance of the system and it is in the time domain that the system must ultimately perform satisfactorily.

A great number of servo design procedures fall within the category of transform techniques. In a broad sense these procedures can be subdivided into a class dominated by a synthesis concept and a class dominated by a compensation concept although in practice neither can be isolated from the other. The synthesis approach to servo design is accomplished in three steps.

1) determination of the closed loop transfer function from specifications
2) determination of the open loop transfer function from the closed loop transfer function obtained in step (1)

3) synthesis of the open loop transfer function with due regard to components which are not at the engineer's disposal.

While the transformed equations are sufficient to determine the system's time domain behavior the complex relationship between the two thwarts ready selection of the "best" transfer function. Hence it is likely that the selected transfer function which meets specifications is more difficult to synthesize than some other transfer function which also meets specifications. Furthermore, there is no systematic or iterative procedure which will improve this initially selected transfer function. To this liability must be added the problems that fixed system components impose on the choice of possible closed loop transfer functions. To its credit the synthesis concept follows a logical sequence from specification to final realization. This is in contrast to the compensation approach which attempts to "patch up" an unsatisfactory initial system in the hope it can eventually be made to meet specifications. The second liability has also been partially eliminated at the expense of simplicity. With the parameters at his disposal the engineer is able to produce the system which most closely approximates the one initially chosen. [1] However this closest point of approach is not guaranteed to meet specifications nor does it preclude that a satisfactory system could be obtained with available parameters were a different choice for the "ideal" closed loop transfer function made.

Utilization of transform techniques to design compensation for an existing skeleton servo-system is an approach more widely employed than synthesis. Two steps are basic to design by compensation.

1) From specifications a "best choice" for the fixed elements of the servo-system is made

2) Then compensation for these elements not under the engineer's control is designed to insure system compliance with specifications.

These steps contrast vividly with those employed in synthesis, largely eliminate the drawbacks of that method and offer very definite advantages. The realization problem is eliminated by the expedient of employing only realizable compensators to modify the fixed skeleton system. Of almost equal importance are the well established methods available for design of compensation. They are for the most part simple, powerful and provide insight in addition to numerical solutions; a feature which is almost unique to these methods. In addition systems designed from a compensation concept tend to be comparatively inexpensive, a fortuitous result which usually occurs when dealing with the physical compensation rather than the theoretical system. However design by compensation also has its disadvantages. Most serious is ignorance of whether the possibility for compensation exists; the fixed system may be incapable of being compensated to meet specifications. In this respect the synthesis concept is clearly superior for whenever an acceptable theoretical system cannot be found then it serves no purpose to attempt design of an impossible physical system.

The preceding thumbnail sketches suffice to acquaint the reader with approaches available to the servo-system designer. However, deficiencies inherent to an approach are subject to the "state of the art" and to the state of development of the methods employed. The compensation approach illustrates the case in point. In an abstract formulation its merits are not outstanding yet the simplicity of the techniques and the well documented correlations between available variables and design criteria serve to make this the most prevalent of approaches. Parameter plane techniques also

advantageously utilize these well established correlations between frequency response, bandwidth, pole-zero patterns and time domain performance specifications such as rise time, maximum overshoot and settling time. Furthermore the essential steady state nature of frequency response analysis and the one parameter limitation of root locus diagrams are largely overcome by parameter plane methods. At the expense of simplicity the effect of simultaneous variation of two parameters on root locations can be displayed. Bandwidth specifications and steady state error requirements can also be indicated on this plot. Furthermore such troublesome questions as second order dominance or whether dominance is possible by adjustment of two parameters can be answered by inspection. These advantages auger more widespread application of parameter plane methods in systems of up to moderate complexity, especially as the digital computer becomes more readily available to compute the required curves.

The intent of this chapter was to introduce the parameter plane and establish its position among tools available to the engineer. The division of servo-system theory into various approaches was arbitrary although it seemed reasonable. Similarly critique of present approaches and methods was pursued as a means of placing the parameter plane and no claim for completeness or comprehensive analysis is made.

CHAPTER II

Concepts of the Parameter Plane

The characteristic equation which is obtained by equating the denominator of the closed loop transfer function to zero, is probably the most important single factor available for determining system performance. In deference to convention the characteristic equation will be designated by

$$f(s) = \sum_{k=0}^{N} a_k s^k = 0 \tag{1}$$

where s is the complex variable of the LaPlace transform

n is the order of the differential equation describing the system

a_k are real constants

However such a designation is an over simplification. Consider the following points.

1. f(s) which is an equation in the complex domain, can properly be regarded as two simultaneous real equations. Namely:

   $$\text{Real Part } \left[f(s)\right] \triangleq f_R(s) = 0$$

   $$\text{Imaginary Part } \left[f(s)\right] \triangleq f_I(s) = 0$$

2. The complex variable s is in reality two real variables since both its real and imaginary parts must be specified before s is fixed. Variables commonly used to specify s are

   $$s = \sigma + j\omega = \omega_N\left[-\zeta + j\sqrt{1-\zeta^2}\right]$$

   For parameter plane formulations the latter specification is the more convenient.

3. Invariably a servo system has a number of parameters which can readily be varied to suit the designer (i.e. system gain) and any such change will affect the coefficients a_k. For parameter plane purposes it is convenient to assume that the a_k are functions of two such variables which are independently at the designer's discretion. That is: $a_k = a_k(\alpha, \beta)$. (The existance of more than two such variables does not invalidate parameter plane concepts any more than the existance of more than one variable invalidated root locus concepts. Rather, all but two of the variable parameters should be fixed prior to any single application of parameter plane technique). With these points in mind a more cogent statement of the characteristic equation would be

$$\begin{aligned} f_R(\alpha, \beta, \zeta, \omega_N) &= \text{Real}\left[\sum_{k=0}^{N} a_k(\alpha,\beta)\, s^k\right] \\ f_I(\alpha, \beta, \zeta, \omega_N) &= \text{Imaginary}\left[\sum_{k=0}^{N} a_k(\alpha,\beta)\, s^k\right] \end{aligned} \qquad (2)$$

Equations (2) are two equations in four unknowns. Classically the coefficients, a_k, were assumed known (the designer assigned values to α, β) and the problem was reduced to one of finding roots for the characteristic equation. If these roots were unsatisfactory the designer would assign new values to α and β and repeat the procedure. This procedure is somewhat illogical in that unknowns (variable system parameters) are initially fixed in expectation of obtaining knowns (desired root locations). A better procedure would be to solve for unknowns in terms of the knowns. That is, solve the implicit equations (2) for explicit expression of α and β.

[illegible]

$$\alpha = \alpha(\zeta, \omega_N)$$

$$\beta = \beta(\zeta, \omega_N) \tag{3}$$

Direct determination of system parameters in terms of desired s-plane locations ζ and ω_N is now feasible. However, s-plane roots are not generally restricted to a single point; rather requirements are such as to place them within a certain region. Suppose that a number of s-plane points on the boundary of this region are chosen, that the corresponding values of system parameters α and β are computed using equations (3) and that these values are plotted as points in a cartesian α-β coordinate system. The curve connecting these points could reasonably be expected to form a boundary for the α-β plane region containing values of system parameters which place roots of the characteristic equation within the desired s-plane region. (A more detailed interpretation of α-β plane curves is reserved for the next chapter). Early investigators were interested in this concept primarily as an aid to stability analysis and therefore restricted their attentions to an s-plane contour along the jw axis [4,5]. The subsequent work of Mitrovic and others generalized procedures to include constant ζ contours but were still limited by restricting α and β to be the last two coefficients of the characteristic equation [3]. More recently Siljak generalized the theory to allow the coefficients of the characteristic equation to contain any linear combination of α and β [7]. This paper extends the procedures further to include the appearance of $\alpha\beta$ products in the coefficients. Still further extension to appearance of α^2 and β^2 terms and beyond is feasible but resulting explicit expressions for $\alpha(\zeta, \omega_N)$ and $\beta(\zeta, \omega_N)$ rapidly become prohibitive.

Parameter plane (or α-β) curves can be considered from an alternative

[illegible]

viewpoint which is more suitable to mathematical rigor. Here (α , β) and (ζ , ω_N) are two sets of variables which are related by either implicit equations (2) or explicit equations (3). The parameter plane (or α-β) curves are then the images (or maps) of the chosen s-plane contours. (The reader is cautioned that the mapping is not conformal and basic relations of conformal mappings are not preserved).

Parameter Plane Equations (Linear Coefficient Case)

This formulation is an abbreviated version of that advanced by Siljak and as such presents only the salient points [7]. Consider the characteristic equation (1) where the coefficients a_k are linear functions of the parameters α and β .

$$f(\alpha , \beta , S) = \sum_{K=0}^{N} (\alpha b_K + \beta c_K + e_K) S^k = 0 \qquad (4)$$

In this equation b_k, c_k and e_k are constants and s is the complex variable of the LaPlace transform. It is convenient to express s as

$$s \triangleq \omega_N \left[-\zeta + j\sqrt{1-\zeta^2} \right] \qquad (5)$$

and

$$s^K \triangleq \omega_N^K \left[T_K(-\zeta) + j\sqrt{1-\zeta^2}\, U_K(-\zeta) \right] \quad K \geq 0 \qquad (6)$$

where w_n is the undamped natural frequency and ζ is the damping constant which assumes positive values in the second quadrant from zero on the imaginary axis to plus one on the negative real axis. Observe that equations (5) and (6) are not inconsistant; rather equation (5) serves to define $T_k(-\zeta)$ and $U_k(-\zeta)$ as explicit functions of ζ .

By manipulation of equations (5) and (6) some subsequently useful relations between the U_k's and T_k's can be obtained.

Argument relations

$$T_k(-\zeta) = (-1)^k \; T_k(\zeta) \tag{7}$$

$$U_k(-\zeta) = (-1)^{k+1} \; U_k(\zeta)$$

Recurrence relations

$$T_{k+1}(\zeta) - 2\zeta T_k(\zeta) + T_{k-1}(\zeta) = 0 \tag{8}$$

$$U_{k+1}(\zeta) - 2\zeta U_k(\zeta) + U_{k-1}(\zeta) = 0$$

Interrelation

$$T_k(\zeta) = \zeta U_k(\zeta) - U_{k-1}(\zeta) \tag{9}$$

Also the $T_k(\zeta)$'s and $U_k(\zeta)$'s are found to be Chebyshev functions of the first and second kind respectively.

Armed with these relations we proceed to find explicit solutions for α and β. Substitution of definition (6) into the characteristic equation as expressed in (4) yields:

$$f(\alpha, \beta, \zeta, \omega_N) = \sum_{k=0}^{N} (\alpha b_k + \beta c_k + e_k)\, \omega_N^k \left[T_k(-\zeta) + j\sqrt{1-\zeta^2}\; U_k(-\zeta)\right] = 0$$

The real and imaginary parts must independently equal zero to satisfy this equation.

$$\sum_{k=0}^{N} (\alpha b_k + \beta c_k + e_k)\, \omega_N^k \, T_k(-\zeta) = 0$$

$$\sum_{k=0}^{N} (\alpha b_k + \beta c_k + e_k)\, \omega_N^k \sqrt{1-\zeta^2}\; U_k(-\zeta) = 0$$

Removing common factors and employing relations (7) to change the arguments of the U_k's and T_k's to positive zeta results in the expressions:

$$\sum_{k=0}^{N} (-1)^k (\alpha b_k + \beta c_k + e_k)\, \omega_N^k \, T_k(\zeta) = 0$$

$$\sum_{k=0}^{N} (-1)^k (\alpha b_k + \beta c_k + e_k)\, \omega_N^k \, U_k(\zeta) = 0$$

Equations (9) which express the interrelation between the T_k's and U_k's can be used to eliminate T_k.

$$\varsigma\left[\sum_{K=0}^{N}(-1)^K(\alpha b_K+\beta c_K+e_K)\,\omega_N^K\,U_K(\varsigma)\right]-\left[\sum_{K=0}^{N}(-1)^K(\alpha b_K+\beta c_K+e_K)\,\omega_N^K\,U_{K-1}(\varsigma)\right]=0$$

$$\left[\sum_{K=0}^{N}(-1)^K(\alpha b_K+\beta c_K+e_K)\,\omega_N^K\,U_K(\varsigma)\right]=0$$

But the quantity in the left hand bracket is identical to the quantity below it which is equal to zero. Incorporating this observation allows us to write

$$\begin{aligned}\sum_{K=0}^{N}(-1)^K\;(\alpha b_K+\beta c_K+e_K)\,\omega_N^K\,U_{K-1}(\varsigma)=0\\ \sum_{K=0}^{N}(-1)^K\;(\alpha b_K+\beta c_K+e_K)\,\omega_N^K\,U_K(\varsigma)=0\end{aligned}\tag{10}$$

Equations (10) are linear in α and β and are easily solved for their explicit expression. This point is readily apparent when equations (10) are rewritten in more concise notation.

$$\begin{aligned}\alpha B_1(\varsigma,\omega_N)+\beta C_1(\varsigma,\omega_N)+E_1(\varsigma,\omega_N)=0\\ \alpha B_2(\varsigma,\omega_N)+\beta C_2(\varsigma,\omega_N)+E_2(\varsigma,\omega_N)=0\end{aligned}\tag{11}$$

where by definition

$$\begin{aligned}B_1(\varsigma,\omega_N)&=\sum_{K=0}^{N}(-1)^K b_K\,\omega_N^K\,U_{K-1}(\varsigma)\\ B_2(\varsigma,\omega_N)&=\sum_{K=0}^{N}(-1)^K b_K\,\omega_N^K\,U_K(\varsigma)\\ C_1(\varsigma,\omega_N)&=\sum_{K=0}^{N}(-1)^K c_K\,\omega_N^K\,U_{K-1}(\varsigma)\\ C_2(\varsigma,\omega_N)&=\sum_{K=0}^{N}(-1)^K c_K\,\omega_N^K\,U_K(\varsigma)\\ E_1(\varsigma,\omega_N)&=\sum_{K=0}^{N}(-1)^K e_K\,\omega_N^K\,U_{K-1}(\varsigma)\\ E_2(\varsigma,\omega_N)&=\sum_{K=0}^{N}(-1)^K e_K\,\omega_N^K\,U_K(\varsigma)\end{aligned}\tag{12}$$

In the interest of brevity the functional dependence of B, C and E will be omitted when such omission can cause no confusion. The explicit solutions for α and β obtained from equations (11) are:

$$\alpha(\zeta, \omega_N) = \frac{C_1 E_2 - C_2 E_1}{B_1 C_2 - B_2 C_1} \tag{13}$$

$$\beta(\zeta, \omega_N) = \frac{B_2 E_1 - B_1 E_2}{B_1 C_2 - B_2 C_1}$$

These are the explicit expressions for α and β which allow mappings of points or contours (except for points on the real axis) from the s-plane into the parameter (or α-β) plane. Applicability does not extend to points on the real axis because the derivation included division by a factor of $(1 - \zeta)^{\frac{1}{2}}$ which is identically zero thereon.

Relevant parameter plane equations for points on the real axis are easily obtained. Define values of the complex variable s on the real axis by

$$s \triangleq -\sigma \tag{14}$$

For values of s on the real axis the characteristic equation (4) becomes

$$f(\alpha, \beta, \sigma) = \sum_{K=0}^{N} (\alpha b_K + \beta c_K + e_K)(-\sigma)^K \tag{15}$$

Of particular importance is the degeneration of the characteristic equation on the real axis into a single real equation whose imaginary part is identically zero. For easier discussion define:

$$B_R(\sigma) = \sum_{K=0}^{N} (-1)^K b_K \sigma^K$$

$$C_R(\sigma) = \sum_{K=0}^{N} (-1)^K c_K \sigma^K \tag{15a}$$

$$E_R(\sigma) = \sum_{K=0}^{N} (-1)^K e_K \sigma^K$$

and write equation (15) as

$$\alpha B_R(\sigma) + \beta C_R(\sigma) + E_R(\sigma) = 0 \tag{16}$$

Equation (16) is a single equation in three unknowns α, β and σ which on selection of a real point (σ_o) to be mapped defines a curve in the α-β plane. For the case under discussion α and β appear linearly and the curve will be a straight line.

Equations (13) and (16) are the keys which allow interpretation of the characteristic equation in the parameter plane. Equation (13) is applicable for all s-plane points except those on the real axis and it maps points into points and s-plane contours into α-β plane contours. Equation (16) represents the degenerate case when s assumes a value on the real axis and it maps a single s-plane point into a straight line in the α-β plane. The significance and interpretation of these parameter curves is explained in detail in the next chapter.

Parameter Plane Equations ($\alpha\beta$ Product Case)

All too often system parameters appear in the characteristic equation non-linearly; an occurence which precludes parameter plane analysis by the techniques thus far developed. Theoretically parameter plane techniques can be extended to cover any eventuality but the impracticability if not the impossibility of obtaining explicit solutions for system parameters in terms of s-plane root locations prevents theory from becoming practice. Characteristic equations in which $\alpha\beta$ products appear are a frequently occuring and important exception to this statement.

$$f(\alpha, \beta, S) = \sum_{K=0}^{N} (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K) S^K = 0 \tag{17}$$

Parameter plane notation and definitions for the $\alpha\beta$ product will be consistent with those employed for the linear case. Noteworthy is the continuing relevance of relations (5) through (9). Separation of equation (17) into its real and imaginary parts yields:

$$\sum_{K=0}^{N} (-1)^K (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K)\, \omega_N^K\, T_K(\zeta) = 0$$

$$\sum_{K=0}^{N} (-1)^K (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K)\, \omega_N^K\, U_K(\zeta) = 0$$

The same manipulations employed in the linear case allow these equations to be expressed as:

$$\sum_{K=0}^{N} (-1)^K (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K)\, \omega_N^K\, U_{K-1}(\zeta) = 0$$
$$\sum_{K=0}^{N} (-1)^K (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K)\, \omega_N^K\, U_K(\zeta) = 0 \qquad (18)$$

In the interests of brevity define:

$$B_1(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K b_K\, \omega_N^K\, U_{K-1}(\zeta)$$
$$B_2(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K b_K\, \omega_N^K\, U_K(\zeta)$$
$$C_1(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K c_K\, \omega_N^K\, U_{K-1}(\zeta)$$
$$C_2(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K c_K\, \omega_N^K\, U_K(\zeta)$$
$$D_1(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K d_K\, \omega_N^K\, U_{K-1}(\zeta) \qquad (19)$$
$$D_2(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K d_K\, \omega_N^K\, U_K(\zeta)$$
$$E_1(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K e_K\, \omega_N^K\, U_{K-1}(\zeta)$$
$$E_2(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K e_K\, \omega_N^K\, U_K(\zeta)$$

and rewrite equations (18) as:

$$\alpha B_1(\zeta, \omega_N) + \beta C_1(\zeta, \omega_N) + \alpha\beta D_1(\zeta, \omega_N) + E_1(\zeta, \omega_N) = 0$$
$$\alpha B_2(\zeta, \omega_N) + \beta C_2(\zeta, \omega_N) + \alpha\beta D_2(\zeta, \omega_N) + E_2(\zeta, \omega_N) = 0 \qquad (20)$$

(Notice that definitions (19) are identical to definitions (12) with $D_1(\zeta, \omega_N)$ and $D_2(\zeta, \omega_N)$ added). Equations (20) can be solved for α and β as functions of ζ and ω_N by three different methods.

1. The parameter β is initially eliminated from equations (20) and α is obtained from the single resulting equation using the quadratic formula. Values for β are subsequently found from a linear formula utilizing the previously found values of α as knowns. Starting with equations (20)

$$\frac{\alpha B_1 - E_1}{\alpha B_2 - E_2} = \frac{\cancel{\beta}(\alpha D_1 - C_1)}{\cancel{\beta}(\alpha D_2 - C_2)}$$

$$\alpha^2(B_1D_2 - B_2D_1) + \alpha(B_1C_2 + D_2E_1 - B_2C_1 - D_1E_2) + (C_2E_1 - C_1E_2) = 0$$

$$\alpha(\zeta, \omega_N) = \frac{(B_2C_1 + D_1E_2 - B_1C_2 - D_2E_1)}{2(B_1D_2 - B_2D_1)} + \frac{\pm\sqrt{(B_2C_1 + D_1E_2 - B_1C_2 - D_2E_1)^2 - 4(B_1D_2 - B_2D_1)(C_2E_1 - C_1E_2)}}{2(B_1D_2 - B_2D_1)} \qquad (21)$$

$$\beta(\zeta, \omega_N) = -\frac{\alpha B_1 + E_1}{\alpha D_1 + C_1} \qquad \beta(\zeta, \omega_N) = -\frac{\alpha B_2 + E_2}{\alpha D_2 + C_2} \qquad (22)$$

The last two formulas yield identical results for β.

2. This method duplicates method 1, except that α is initially eliminated and β is solved via the quadratic formula.

$$\beta(\zeta, \omega_N) = \frac{(B_1C_2 + D_1E_2 - B_2C_1 - D_2E_1)}{2(C_1D_2 - C_2D_1)} + \frac{\pm\sqrt{(B_2C_1 + D_2E_1 - B_1C_2 - D_1E_2)^2 - 4(C_1D_2 - C_2D_1)(B_2E_1 - B_1E_2)}}{2(C_1D_2 - C_2D_1)} \qquad (23)$$

$$\alpha(\zeta, \omega_N) = -\frac{\beta C_1 + E_1}{\beta D_1 + B_1} \qquad \alpha(\zeta, \omega_N) = -\frac{\beta C_2 + E_2}{\beta D_2 + B_2} \qquad (24)$$

3. A third procedure involves solution for both $\alpha(\zeta, \omega_N)$ and $\beta(\zeta, \omega_N)$ by the quadratic formula after β and α,

respectively are eliminated among equations (20). The results are solutions (21) and (23) which will be designated

$$\alpha(\zeta, \omega_N) = X \pm \sqrt{Y}$$
$$\beta(\zeta, \omega_N) = U \pm \sqrt{V}$$

While equations (21) through (24) display explicit solutions for α and β a relevant question still remains - are solutions (21) and (22) identical to those obtained in (23) and (24)? This question paraphrased to apply to the third procedure would be - are each of the possible point sets (α, β) valid solutions for equations (20) or are two of these sets extraneous? Direct substitution reveals that the solutions obtained using equations (21) and (22) are identical to the solutions obtained using equations (23) and (24). Moreover, using the third procedure yields valid solutions:

$$(\alpha, \beta) = (x + \sqrt{y},\ u - \sqrt{v}) \quad \text{and} \quad (\alpha, \beta) = (x - \sqrt{y},\ u + \sqrt{v})$$

The two remaining possibilities are extraneous and will not satisfy equations (20).

As for the linear case s-plane points on the real axis are degenerate; the imaginary restriction on the characteristic equation is identically satisfied and only one equation relates the three variables α, β and s. This equation which is appropriate for points on the real axis is obtained by defining s as in (14) and substituting into characteristic equation (17).

$$f(\alpha, \beta, \sigma) = \sum_{K=0}^{N} (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K)(-\sigma)^K = 0 \qquad (25)$$

As before define:

$$B_R(\sigma) = \sum_{K=0}^{N} (-1)^K b_K \sigma^K$$
$$C_R(\sigma) = \sum_{K=0} (-1)^K c_K \sigma^K \qquad (25a)$$

$$D_R(\sigma) = \sum_{k=0}^{N} (-1)^k d_k \sigma^k$$
$$E_R(\sigma) = \sum_{k=0} (-1)^k e_k \sigma^k$$

and write equation (25) as:

$$\beta(\alpha, \sigma) = -\frac{\alpha B_R(\sigma) + E_R(\sigma)}{\alpha D_R(\sigma) + C_R(\sigma)} \qquad (26)$$

The locus of this equation for any given value of σ describes a hyperbola in the parameter plane with asymptotes of:

$$\alpha = -\frac{C_R(\sigma)}{D_R(\sigma)} \quad \text{and} \quad \beta = -\frac{B_R(\sigma)}{D_R(\sigma)}$$

Comparing these explicit expressions with those occuring for the case of coefficients linear in α and β we find three mentionable differences.

1. In the linear case all s-plane values mapped into real values in the α-β plane. For the $\alpha\beta$ product case this is no longer true; the possibility for complex α, β values exists. Physically this implies that it is impossible to place s-plane roots at such locations by adjustment of α and β.
2. In the linear case each complex s-plane point mapped into a single point in the parameter plane. For $\alpha\beta$ products a complex s-plane point usually maps into two different points on the α-β plane.
3. For both linear and product cases real s-plane points have curves as α-β plane images; however, where the curve is a straight line for the linear case it becomes a hyperbola for the product case.

Parameter Plane Methods Extended Beyond the Product Case

Extension of parameter plane methods beyond the product case appears prohibitive even when explicit solutions for α and β can be found.

Suppose we allow characteristic equations of the form

$$f(\alpha, \beta, S) = \sum_{k=0}^{N} \left(\alpha^2 a_{1k} + \alpha\beta a_{2k} + \beta^2 a_{3k} + \alpha a_{4k} + \beta a_{5k} + a_{6k}\right) S^k = 0$$

Under these assumptions explicit solutions for α and β exist and involve no insurmountable or extraordinary difficulties; in fact simultaneous solution of two conics for α and β yields the desired expressions.

$$\alpha^2 A_{11} + \alpha\beta A_{21} + \beta^2 A_{31} + \alpha A_{41} + \beta A_{51} + A_{61} = 0$$
$$\alpha^2 A_{12} + \alpha\beta A_{22} + \beta^2 A_{32} + \alpha A_{42} + \beta A_{52} + A_{62} = 0$$

where

$$A_{11}(\zeta, \omega_N) = \sum_{k=0}^{N} (-1)^k a_{1k} \omega_N^k U_{k-1}(\zeta)$$
$$A_{32}(\zeta, \omega_N) = \sum_{k=0}^{N} (-1)^k a_{3k} \omega_N^k U_k(\zeta)$$

The problem of consequence is interpretation of the parameter plane curves since each s-plane point generally has four α-β plane images. Additionally, the four valid solutions for (α, β) must be distinguished from among a minimum of eight generated in the process of solving the simultaneous equations.

Parameter Plane Methods for More Than Two Parameters

The parameter plane is a competent two variable technique in analogy to the root locus being a capable single variable technique. Pursuing this analogy we find that parameter plane technique can accommodate three variables with a facility about equal to that of the root locus to accommodate two variables. Problems having a third parameter plane variable can be approached in three ways.

1. An additional restriction or specification is invoked (such as a steady state error restriction) which permits solution

for the third variable, γ , as a function of the remaining two, f(α , β). With elimination of the third parameter the problem reverts to the standard two parameter problem and analysis proceeds as before. However linearity of α and β in the coefficients of the characteristic equation is commonly destroyed when eliminating γ in favor of f(α , β). This shortcoming seriously diminishes this otherwise attractive approach since loss of linearity usually precludes parameter plane analysis.

2. Addition of a third dimension to the two dimensional parameter plane allows adequate representation of the three variable problem. Under this representation an s-plane contour maps into a surface in (α , β , γ) parameter space. Any choice of system parameters which lie on this surface will have at least one pair of roots on the chosen s-plane contour. Only the problems innate to three dimensional presentation limit the attractiveness of this method.

3. One selected s-plane contour can be mapped into several α-β plane curves, each curve representing a different value of γ, the third parameter. This procedure effectively extends parameter plane analysis to three variables especially when only one or two s-plane contours are of interest. Another credit is that applicability is not limited by the manner in which the third parameter appears in the characteristic equation. Moreover, the α-β plane image of any specified s-plane point (real axis points excluded) plots on a straight line as the third parameter is varied whenever all parameters are linear

with respect to one another.

This last observation is quite useful in the following context. Suppose we specify any s-plane point as s(x,y) where x is held constant and y is the running variable along the s-plane contour of interest. The map of this contour in the α-β plane depends on the third parameter; a different mapping for each different value of γ . Connect the points which, on each curve, represent the same value of the running parameter y - the locus of these points is a straight line in the parameter plane. Furthermore, the spacing of the various image curves is proportional to the value of γ .

The apparent utility of these linearity properties warrants the more rigorous foundation supplied by formal theorem and proof.

<u>Theorem I</u>. Suppose three variable parameters appear linearly in the coefficients of the characteristic equation.

$$f(\alpha, \beta, \gamma, s) = \sum_{K=0}^{N} (\alpha b_K + \beta c_K + \gamma f_K + e_K) s^K = 0 \tag{27}$$

Then the locus of any fixed s-plane point in the parameter ($\alpha - \beta$) plane is a straight line and the image of the point is a linear function of the third parameter .

<u>Proof</u>. Resolve equation (27) into its real and imaginary parts, equate each part to zero and simplify as was done when originally formulating the parameter plane equations. The result is two equations analogous to equations (11).

$$\begin{aligned} &\alpha B_1(\zeta, \omega_N) + \beta C_1(\zeta, \omega_N) + \gamma F_1(\zeta, \omega_N) + E_1(\zeta, \omega_N) = 0 \\ &\alpha B_2(\zeta, \omega_N) + \beta B_2(\zeta, \omega_N) + \gamma F_2(\zeta, \omega_N) + E_2(\zeta, \omega_N) = 0 \end{aligned} \tag{28}$$

where

$$F_1(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K f_K \omega_N^K U_{K-1}(\zeta) \tag{28a}$$

$$F_2(\zeta, \omega_N) = \sum_{K=0}^{N} (-1)^K f_K \omega_N^K U_K(\zeta)$$

and all other quantities are as defined in (12)

Explicit solution of equations (28) yields the customary parameter plane equations.

$$\alpha(\gamma, \zeta, \omega_N) = \frac{\gamma(C_1F_2 - C_2F_1) + (C_1E_2 - C_2E_1)}{(B_1C_2 - B_2C_1)}$$
$$\beta(\gamma, \zeta, \omega_N) = \frac{\gamma(B_2F_1 - B_1F_2) + (B_2E_1 - B_1E_2)}{(B_1C_2 - B_2C_1)} \tag{29}$$

For a fixed point in the s-plane ζ and ω_N become knowns and α and β become functions of the single variable γ. Under these conditions equations (29) are recognizable as the parametric representation of a straight line in the $\alpha-\beta$ plane. Furthermore the linear appearance of γ in both of these explicit expressions implies that the position of the image point on the straight line locus is linearly proportional to the value of γ. Q.E.D.

Parameter Plane Equations in the Z-Plane

The aforementioned analogy between parameter plane methods in the s and z-planes can be strengthened to practical identity of procedures and equations. (The only noteworthy difference between the two planes is choice of contours to be mapped onto the $\alpha-\beta$ plane; constant ζ contours are popular in the s-plane whereas spirals of constant damping or circles of constant settling time are popular in the z-plane). This close correspondence made formulation of the parameter plane equations in the more familiar s-plane with subsequent extension to the z-plane the most attractive alternative. For z-plane formulation we assume a characteristic equation of the form

$$f(\alpha, \beta, z) = \sum_{K=0}^{N} (\alpha b_K + \beta c_K + e_K) Z^K = 0$$

with provision for inclusion of terms for the $\alpha\beta$ product case, three parameter case, etc. Results derived in the s-plane are valid in the z-plane subject to the following modifications.

a) If the complex variable z is defined as

$$z = \omega_{N_z}\left(-\zeta_z + j\sqrt{1-\zeta_z^2}\right)$$

With z so defined no modification of s-plane formulas is necessary beyond replacing ζ by ζ_z and ω_N by ω_{N_z}.

b) If the complex variable z is defined as

$$z = \omega_{N_z}\left(\zeta_z + j\sqrt{1-\zeta_z^2}\right)$$

The total effect of so defining z is to obviate the need for the $(-1)^K$ factor in the definitions of B,C,D,E and F. Specifically, all s-plane formulas are valid if the definitions in (12), (15a), (19) (25a) and (28) are redefined as

$$\begin{aligned}
B_1(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} b_K \omega_{N_z}^K U_{K-1}(\zeta_z) \\
B_2(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} b_K \omega_{N_z}^K U_K(\zeta_z) \\
C_1(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} c_K \omega_{N_z}^K U_{K-1}(\zeta_z) \\
C_2(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} c_K \omega_{N_z}^K U_K(\zeta_z) \\
D_1(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} d_K \omega_{N_z}^K U_{K-1}(\zeta_z) \\
D_2(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} d_K \omega_{N_z}^K U_K(\zeta_z) \\
E_1(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} e_K \omega_{N_z}^K U_{K-1}(\zeta_z) \qquad (30) \\
E_2(\zeta_z, \omega_{N_z}) &= \sum_{K=0}^{N} e_K \omega_{N_z}^K U_K(\zeta_z)
\end{aligned}$$

$$F_1(\zeta_z, \omega_{N_z}) = \sum_{K=0}^{N} f_K \, \omega_{N_z}^K \, U_{K-1}(\zeta_z)$$

$$F_2(\zeta_z, \omega_{N_z}) = \sum_{K=0}^{N} f_K \, \omega_{N_z}^K \, U_K(\zeta_z)$$

$$B_R(\sigma_z) = \sum_{K=0}^{N} b_K \, \sigma_z^K$$

$$C_R(\sigma_z) = \sum_{K=0}^{N} c_K \, \sigma_z^K$$

$$D_R(\sigma_z) = \sum_{K=0}^{N} d_K \, \sigma_z^K$$

$$E_R(\sigma_z) = \sum_{K=0}^{N} e_K \, \sigma_z^K$$

$$F_R(\sigma_z) = \sum_{K=0}^{N} f_K \, \sigma_z^K$$

The reader is cautioned that the analogy between ζ and ζ_z and ω_N and ω_{N_z} extends only to the mathematical formulation of the parameter plane equations. The value of ζ_z is in no way indicative of system overshoot nor is ω_{N_z} indicative of the system's undamped natural frequency.

Mitrovic's Method and Extensions of Mitrovic's Method

Although Mitrovic methods are special cases of the parameter plane they warrant mention because of a relation between the $\alpha' - \beta'$ (M-plane) curves and the s-plane roots not present in the more general case. This relation and its application to the z-plane will be discussed in Chapter 3; the present objective is simply to define Mitrovic methods. Parameter plane theory assumes a characteristic equation of the form:

$$f(\alpha, \beta, S) = \sum_{K=0}^{N} (\alpha \, b_K + \beta \, c_K + e_K) \, S^K = 0$$

Mitrovic's method and its extensions assume that only one b_i and one c_j ($i \neq j$) have non-zero value; α and β are subsequently redefined as:

$$\alpha' = \alpha b_i$$

$$\beta' = \beta c_j$$

The characteristic equation in a format apropos to Mitrovic formulation then appears as:

$$f(\alpha', \beta', s) = \sum_{\substack{k=0 \\ k \neq i, j}}^{N} e_k s^k + \alpha' s^i + \beta' s^j = 0$$

If i = 0, j = 1 the formulation is the one originally advocated by Mitrovic; otherwise the formulation is referred to as an extension of Mitrovic's method.

CHAPTER III

The Interpretation Problem

To regard a functional relation between two complex variables, w = g(z), as limited to defining w in terms of z is indeed tunnel vision. Lost would be such concepts as images of z-plane contours and mappings of z-plane areas into selected areas of the w-plane. An equally grievous error would be to regard the explicit expressions for $\alpha(\zeta, \omega_N)$ and $\beta(\zeta, \omega_N)$ as formulas to place roots at specified s-plane (z-plane) locations. Such regard would obscure two most important interpretations, that of mapping and that of partition of the $\alpha-\beta$ plane into regions having a similar root property. The problem, therefore, is not for a single interpretation of $\alpha-\beta$ plane curves but rather it is for an answer to the question, "How can the significant information buried in s-plane (z-plane) pole zero configurations be extracted from parameter plane curves?"

Interpretation in the $\alpha-\beta$ Plane

A simple yet comprehensive answer to the interpretation problem is provided by examining various point sets in the $\alpha-\beta$ plane for common properties. In figure (1a) a z-plane contour terminating at two points on the real z-axis is designated; figure (1b) shows the $\alpha-\beta$ plane image of this contour and the images of the two real z-plane points terminating this contour. With reference to figure (1b) define the following point sets.

Point Set A - Point set A contains the images of all points on the designated z-plane contour.

Note: This set contains all points on the image curve drawn in figure (1b) and no points not on this curve.

Point Set B - Point set B contains all points which are the image of the real z-plane points σ_{z_1} and σ_{z_2} .

Point Set C - Point set C contains all points in the cross-hatched area and it is bounded by but disjoint from set A.

Point Set D - Point set D contains all points in the dashed shaded area and it is disjoint from set A and set B.

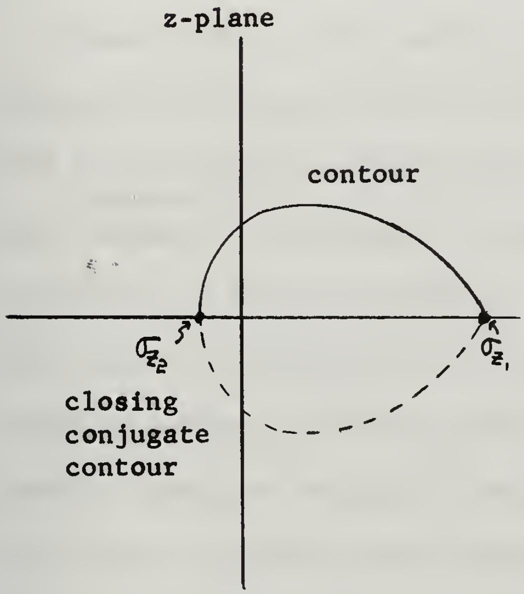


FIGURE 1a

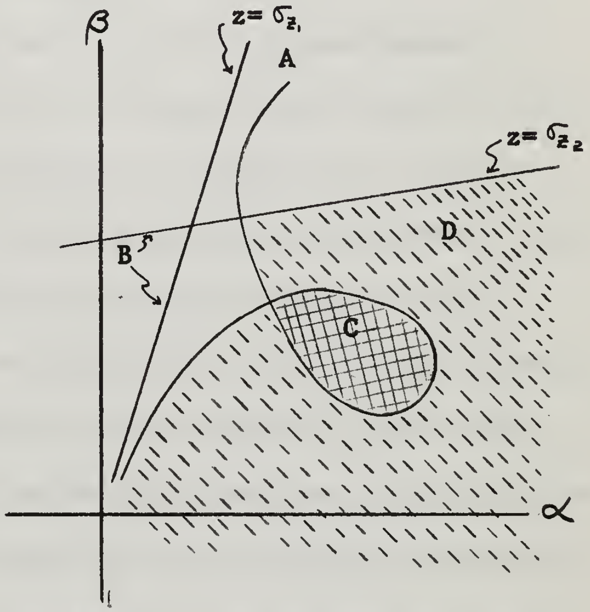


FIGURE 1b

Let M(α , β) be any point in the parameter plane and consider the significance of M belonging to certain point sets.

Observation: If M belongs to set A the characteristic equation must have one complex root lying on the chosen z-plane contour and another root lying on its reflection in the real axis. Recall that explicit equations

$$\alpha = \alpha(\zeta_z, \omega_{N_z})$$
$$\beta = \beta(\zeta_z, \omega_{N_z})$$

give the parameter values needed to place a root at (ζ_z , ω_{N_z}). However, these explicit equations can also be considered as mapping functions which establish a relation between z-plane points and $\alpha-\beta$ points. Regarded thus, every point in set A has its domain on the designated z-plane contour and satisfies these explicit expressions; hence the conclusion observed is inevitable.

Observation: If M is contained in set B the characteristic equation must have at least one real root at $z = \sigma_{z_1}$ (or σ_{z_2}).

Observation: All M-points located in set C have an equal number of characteristic equation roots enclosed by the z-plane contour and its reflection in the real axis. Furthermore the number of characteristic equation roots enclosed by an M-point in set C differs by two from the number of roots enclosed by an M-point in set D. The validity of this statement can be established by contradiction. Assume two points exist in set C which have a different number of characteristic equation roots enclosed by the z-plane contour. Then, by continuous motion of the M-point from one of these points to the other, characteristic equation roots could move across the closed z-plane contour without ever once having the roots lie on the contour itself. (All M-points having roots on the contour belong to set A which is disjoint from set C.) This is impossible; therefore by contradiction all points in C have the same number of characteristic equation roots enclosed by the z-plane contour. Similar continuity arguments can be applied to show that a different number of roots are enclosed for M-points in set C and set D. (To deny this proposition would imply that a root lying on the specified z-contour could be moved in any direction in the z-plane without affecting the number of roots enclosed by the contour). That the enclosed roots in set C and D differ by two stems from the restriction that coefficients of the characteristic equation be real; root motion across the contour's upper limb requires motion of

the conjugate root across the contour's lower limb.

Observation: The $\alpha-\beta$ plane images of points at which the z-plane contour intersects the real z-axis also serve to divide the parameter plane into point sets. M-points belonging to sets on opposite sides of such contours will differ by one in the number of characteristic equation roots enclosed by the z-plane contour. Justification of this statement is analogous to that of the preceding observation.

Preceding observations although they cannot claim mathematical rigor provide a framework for parameter plane interpretation one step above the intuitive approach usually advanced. For application these observations can be profitably restated.

1. M-points on an image curve imply a characteristic equation root on the z-plane (s-plane) contour upon whose image the M-point lies
2. Parameter plane images of z-plane contours and images of the intersection points of these contours with the real axis serve to divide the $\alpha-\beta$ plane into regions or point sets. The distinguishing characteristic of each region is that all M-points within a region have the same number of characteristic equation roots enclosed by the z-plane (s-plane) contour.
3. If the boundary between regions is the image of a complex contour the number of roots enclosed for M-points located in adjacent regions differs by two. If the boundary between regions is the image of a real axis point the number of roots enclosed for M-points in adjacent regions differs by one.

Shading Convention and Rules

The interpretation of $\alpha-\beta$ curves as images of selected s-plane (z-plane) countours was previously discussed. Moreover, the preceding paragraph showed that movement of the M-point across an $\alpha-\beta$ curve corresponds to two complex roots moving across the selected s-plane (z-plane) contour. The direction in which the roots move across the s plane contour is usually available by inspection of the direction in which the M-point moves across the $\alpha-\beta$ curve. (i.e., Motion of the M-point across the $\alpha-\beta$ plane image of the contour $\zeta=\zeta_0$ in the direction of $\alpha-\beta$ plane images of contours $\zeta > \zeta_0$ implies that the s-plane roots move across the contour in the direction of increasing ζ). However, parameter plane regions in which images of two different ζ contours intersect or in which the image of a given s-plane contour loops on itself do not readily allow the direction of root motion across the s-plane contour to be inferred from the motion of the M-point across its $\alpha-\beta$ plane image. For such cases, and in fact whenever required, the direction of root motion across an s-plane contour as the M-point moves across its image in the parameter plane can be determined mathematically.

Adopt the convention that the parameter plane images of constant ζ (or constant ω_N) s-plane contours are shaded so that an M-point crossing the $\alpha-\beta$ curve from shaded to unshaded side causes two complex s-plane roots to leave that portion of the s-plane enclosed by the contour.

Theorem I The sign of the Jacobian $J\left(\frac{\alpha,\beta}{\omega_N\ \zeta}\right)$ determines the side of the image curve to be shaded. If the Jacobian is positive (negative) the left (right) hand side of the image curve should be shaded facing in the direction of increasing ω_N (ζ).

Proof Construct a cartesian three dimensional vector space on the

parameter plane whose position vector is:

$$\vec{r}(\alpha, \beta, z) = \alpha \vec{i} + \beta \vec{j} + z \vec{k}$$

In this space all curves lying in the parameter plane, as all images of s-plane contours must, will have position vectors of the form

$$\vec{r}(\alpha, \beta) = \alpha \vec{i} + \beta \vec{j}$$

Now recall that a requirement for the construction of parameter plane curves was the existence of explicit solutions for α and β in terms of ζ and ω_N. Namely:

$$\alpha = \alpha(\zeta, \omega_N)$$
$$\beta = \beta(\zeta, \omega_N)$$

This requirement allows the position vector $\vec{r}(\alpha, \beta)$ to be written as

$$\vec{r}(\zeta, \omega_N) = \alpha(\zeta, \omega_N)\vec{i} + \beta(\zeta, \omega_N)\vec{j}$$

The $\alpha-\beta$ plane image of a constant ζ contour of value ζ_0 is traced by the vector

$$\bar{r}(\zeta_0, \omega_N) = \alpha(\zeta_0, \omega_N)\vec{i} + \beta(\zeta_0, \omega_N)\vec{j}$$

Fix attention at one point on this curve and determine the motion of that point resulting from an incremental increase in ζ. It is

$$\overrightarrow{dr}(\zeta_0 + \Delta\zeta, \omega_0) = \frac{\overrightarrow{\partial r}(\zeta_0, \omega_0)}{\partial \zeta}\, d\zeta$$

Obviously the direction of M-point motion in the direction of increasing ζ at this point is determined by the vector $\frac{\overrightarrow{\partial r}(\zeta_0, \omega_0)}{\partial \zeta}$. Since an M-point moving in a direction to increase ζ causes roots to enter the area enclosed by a constant ζ contour that side of the $\alpha-\beta$ plane image toward

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

which the vector $\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \varsigma}$ points is shaded. Graphically this scheme suffices but it can also be stated in a mathematical formulation. By definition the tangent to the α-β plane image of the constant ς_o s-plane contour at the point (ς_o , ω_o) in the direction of increasing ω_N is given by the vector $\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \omega_N}$. If the counter-clockwise rotation of the vector $\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \omega_N}$ into the vector $\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \varsigma}$ is more than 180° the latter vector points to the right of the constant ς_o curve image and that side of the α-β curve should be shaded. Contrariwise, if the rotation is less than 180° the vector $\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \varsigma}$ points to the left of the constant ς_o mapping in the α-β plane and that side of the curve should be shaded. Fortunately this criterion is the exact definition of the cross-product of two vectors lying in a cartesian plane.

$$\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \omega_N} \times \frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \varsigma} = A\vec{k}$$

If A is negative shade the right hand side of the α-β plane curve facing in the direction of increasing ω_N ; if it is positive shade the left hand side. At this point reconcile the sign of A with a more familiar expression. To do this express $\vec{r}(\varsigma, \omega_N)$ in terms of its α and β components and take the required derivatives.

$$\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \omega_N} = \frac{\partial \alpha(\varsigma_o, \omega_o)}{\partial \omega_N}\vec{i} + \frac{\partial \beta(\varsigma_o, \omega_o)}{\partial \omega_N}\vec{j}$$

$$\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \varsigma} = \frac{\partial \alpha(\varsigma_o, \omega_o)}{\partial \varsigma}\vec{i} + \frac{\partial \beta(\varsigma_o, \omega_o)}{\partial \varsigma}\vec{j}$$

and

$$\frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \omega_N} \times \frac{\partial \vec{r}(\varsigma_o, \omega_o)}{\partial \varsigma} = \begin{vmatrix} \frac{\partial \alpha(\varsigma_o, \omega_o)}{\partial \omega_N} & \frac{\partial \beta(\varsigma_o, \omega_o)}{\partial \omega_N} \\ \frac{\partial \alpha(\varsigma_o, \omega_o)}{\partial \varsigma} & \frac{\partial \beta(\varsigma_o, \omega_o)}{\partial \varsigma} \end{vmatrix} \vec{k}$$

Hence the sign of A depends on the sign of the determinant which is by definition the well known Jacobian $J\left(\frac{\alpha\ \beta}{\omega_N\ \zeta}\right)$.

Extension of these results to other s-plane contours and also to z-plane contours is possible. In fact some general theorems must exist which encompasses all alternatives, but the statement of such a theorem and the determination of necessary and sufficient conditions are difficult. Since there are but a limited number of commonly used s-plane (z-plane) contours it suffices to list the governing Jacobian for each. The proof in each case proceeds exactly as that presented above.

<u>Extension of Theorem I</u> The sign of the Jacobian $J\left(\frac{\alpha, \beta}{x\ y}\right)$, where the independent variables x,y are dependent on the contour being mapped, determines the side of the image curve in the α-β plane to be shaded. Shading rules are as listed in theorem I.

a) For constant settling time or frequency contours in the s-plane the governing Jacobian is

$$J\left(\frac{\alpha\ \beta}{\omega\ \sigma}\right) = \begin{vmatrix} \frac{\partial\alpha}{\partial\omega_N}\frac{\partial\omega_N}{\partial\omega} + \frac{\partial\alpha}{\partial\zeta}\frac{\partial\zeta}{\partial\omega} & \frac{\partial\beta}{\partial\omega_N}\frac{\partial\omega_N}{\partial\omega} + \frac{\partial\beta}{\partial\zeta}\frac{\partial\zeta}{\partial\omega} \\ \frac{\partial\alpha}{\partial\omega_N}\frac{\partial\omega_N}{\partial\sigma} + \frac{\partial\alpha}{\partial\zeta}\frac{\partial\zeta}{\partial\sigma} & \frac{\partial\beta}{\partial\omega_N}\frac{\partial\omega_N}{\partial\sigma} + \frac{\partial\beta}{\partial\zeta}\frac{\partial\zeta}{\partial\sigma} \end{vmatrix}$$

where $s = \omega_N(-\zeta + j\sqrt{1-\zeta^2}) = -\sigma + j\omega$

b) For constant ζ_z or ω_{N_z} (settling time) contours in the z-plane the governing Jacobian is

$$J\left(\frac{\alpha\ \beta}{\omega_{N_z}\ \zeta_z}\right) = \begin{vmatrix} \frac{\partial\alpha}{\partial\omega_{N_z}} & \frac{\partial\beta}{\partial\omega_{N_z}} \\ \frac{\partial\alpha}{\partial\zeta_z} & \frac{\partial\beta}{\partial\zeta_z} \end{vmatrix}$$

where $z = \omega_{N_z}(-\zeta + j\sqrt{1-\zeta^2})$

Note: Frequently z is expressed as z = $\omega_{N_z}\left(\zeta_z + j\sqrt{1-\zeta_z^2}\right)$ in which case the sign of the Jacobisn is reversed.

c) For constant damping contours in the z-plane the governing Jacobian is

$$J\left(\frac{\alpha,\beta}{\omega_N,\zeta}\right) = \begin{vmatrix} \frac{\partial\alpha}{\partial\omega_{N_z}}\frac{\partial\omega_{N_z}}{\partial\omega_N} + \frac{\partial\alpha}{\partial\zeta_z}\frac{\partial\zeta_z}{\partial\omega_N} & \frac{\partial\beta}{\partial\omega_{N_z}}\frac{\partial\omega_{N_z}}{\partial\omega_N} + \frac{\partial\beta}{\partial\zeta_z}\frac{\partial\zeta_z}{\partial\omega_N} \\ \frac{\partial\alpha}{\partial\omega_{N_z}}\frac{\partial\omega_{N_z}}{\partial\zeta} + \frac{\partial\alpha}{\partial\zeta_z}\frac{\partial\zeta_z}{\partial\zeta} & \frac{\partial\beta}{\partial\omega_{N_z}}\frac{\partial\omega_{N_z}}{\partial\zeta} + \frac{\partial\beta}{\partial\zeta_z}\frac{\partial\zeta_z}{\partial\zeta} \end{vmatrix} .$$

Mathematical formulation for shading is virtually complete except for the degenerate case of points on the real axis. Each of these points have an entire curve as their image in the parameter plane since only one equation exists (16) which relates the three variables α, β, σ. (The restriction that the imaginary part of the characteristic equation equal zero is identically satisfied by points on the real axis). If parameter plane curves can be constructed explicit solution of equation (16) for α in terms of β and σ or for β in terms of α and σ exists. Designate these explicit expressions as

$$\begin{aligned} \alpha &= \alpha(\beta,\sigma) \\ \beta &= \beta(\alpha,\sigma) \end{aligned} \tag{31}$$

<u>Theorem II</u> The shading of the parameter plane curve which is the image of a real point σ(σ_z) in the s-plane (z-plane) is determined by the sign of the partial derivative $\frac{\partial\beta(\alpha_0,\sigma_0)}{\partial\sigma}$ (or $\frac{\partial\alpha(\beta_0,\sigma_0)}{\partial\sigma}$) where (α_0 , β_0) is a point on the parameter plane image of σ_0 . The shading rule is as follows.

(a) If increasing σ (or σ_z) from σ_0 causes a real root to enter the area enclosed by the contour shade the left (right) side of the curve facing in the direction of increasing α when $\frac{\partial\beta(\alpha_0,\sigma_0)}{\partial\sigma}$ is positive (negative).

Equivalent to this is the statement that if increasing σ causes a real root to enter the area enclosed by the contour, shade the left (right) hand side of the curve facing in the direction of increasing β when $\frac{\partial\alpha(\beta_0,\sigma_0)}{\partial\sigma}$ is negative (positive).

(b) If increasing σ (or σ_z) from σ_0 causes a real root to leave the area enclosed by the s-plane (or z-plane) contour shade the right (left) side of the α-β curve facing in the direction of increasing α when $\frac{\partial\beta(\alpha_0,\sigma_0)}{\partial\sigma}$ is positive (negative).

For comparison a table presents this theorem in a most effective manner.

Case	Face Direction	Sign of $\frac{\partial\beta(\alpha_0,\sigma_0)}{\partial\sigma}$	Shade	Face Direction	Sign of $\frac{\partial\alpha(\beta_0,\sigma_0)}{\partial\sigma}$	Shade
Increased σ (or σ_z) implies a root enters the enclosed area	Increasing α	Positive	Left	Increasing β	Negative	Left
		Negative	Right		Positive	Right
Increased σ (or σ_z) implies a root leaves the enclosed area	Increasing α	Positive	Right	Increasing β	Negative	Right
		Negative	Left		Positive	Left

Table I

Values on the real s-axis and the real z-axis are not defined identically; the usual case being

$$\text{Real } s = -\sigma$$

$$\text{Real } z = +\sigma_z$$

The manner of definition does not affect the validity of the theorem if the user is consistent throughout. For instance, under these definitions a change of s from -1 to -2 would indicate an increase in σ whereas a similar change in z would indicate as a decrease in σ_z . Notice also that either $\frac{\partial\beta(\alpha_0,\sigma_0)}{\partial\sigma}$ or $\frac{\partial\alpha(\beta_0,\sigma_0)}{\partial\sigma}$ carry sufficient information to shade and their dictates must always agree. An exception occurs when the tangent to the α-β curve is vertical or horizontal in which case one of the partials will degenerate to a value of zero or infinity while the other will still impart useful information.

The proof to this theorem is analogous to that presented for theorem I. In this case either α and σ or β and σ can be considered independent variables depending on which variable (α or β) is explicitly expressed. The position vectors are

$$\vec{r}(\alpha,\sigma) = \alpha\vec{i} + \beta(\sigma,\alpha)\vec{j} \quad \text{for explicit expression of } \beta$$
$$\vec{r}(\beta,\sigma) = \alpha(\beta,\sigma)\vec{i} + \beta\vec{j} \quad \text{for explicit expression of } \alpha$$

From this point forward the proof is identical with that of theorem I.

A few concluding comments on shading and shading formulas should be reiterated for the reader. First, the shading of parameter plane curves provides a comprehensive picture of root motion across selected s-plane contours as a function of the simultaneous variation of two system parameters. When several s-plane contours are mapped onto the α-β plane this root motion is usually evident without the tedium of shading and if shading is desired it can be accomplished by inspection. For any case where inspection shading is impossible the formulas presented here resolve the difficulty.

In this they extend beyond existing criteria which can be applied only for coefficients a_k linear in α and β [4]. However, the labor involved in evaluating the sign of a Jacobian is somewhat greater than that required for existing shading procedures and the former undoubtedly will find employment only as a last resort. Probably the most significant contribution of these rules is that they put shading on a more rigorous framework.

Example of Parameter Plane Curve Interpretation.

An example will serve to illustrate the salient features of parameter plane interpretation and curve shading discussed in the preceding sections. Consider the third-order error-sampled system with zero-order-hold and tachometer feedback shown in figure 2. Suppose that there is a requirement that roots for the sampled transfer function have zeta values greater than 0.5 or lie within the corresponding spiral contour in the z-plane. Allowable s-plane and z-plane root areas are shaded in figures 3a and 3b respectively. Proper adjustments of k and k_t to meet this requirement are available from interpretation of the parameter plane image of the z-plane contour which bounds the allowable roots. Choose α and β as k and kk_t respectively; for this choice the characteristic equation of the system shown in figure 2 is:

$$f(\alpha, \beta, z) = z^3 + (.058\alpha + .310\beta - 1.974)z^2 + \qquad (32)$$

$$(.163\alpha - .122\beta + 1.198)z + (.028\alpha - .188\beta - .223) = 0$$

The coefficients of this characteristic equation are linear in α and β; α-β plane curves (mappings) are obtained by utilizing the parameter plane formulas for the linear coefficient case developed in chapter 2. These image curves and the images of the two points at which the designated z-plane contour intersects the real axis are shown in figure 4. Shading in this figure is by inspection. (In this case inspection shading can be

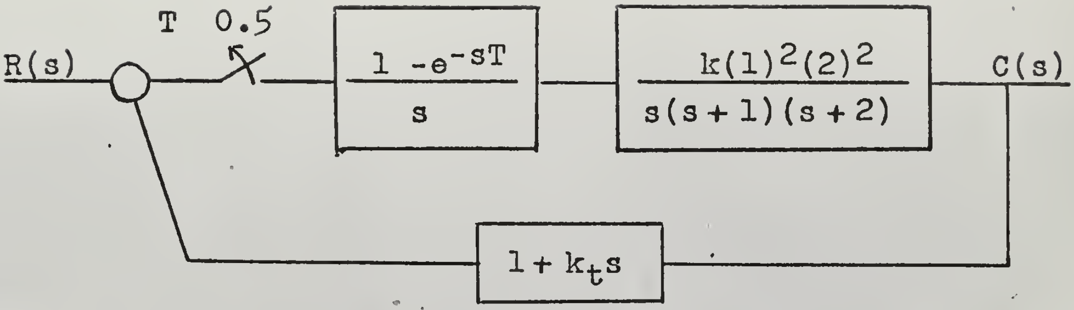
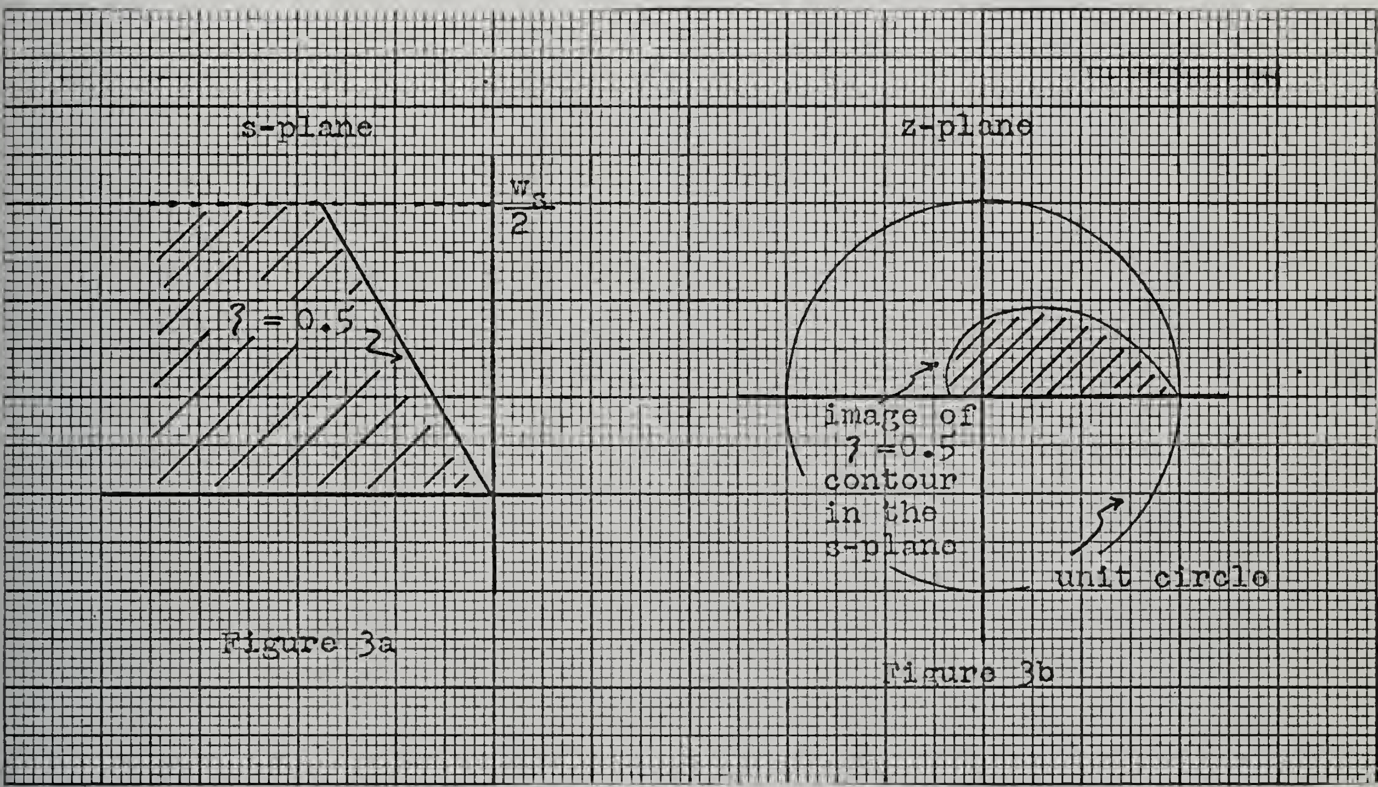


Figure 2



M-Point	Values Alpha	Beta	Characteristic Equation Roots		Number of Roots Within Contour
M1	.4	.5	.542 ± j369	.712	3
M2	.4	.1	.808 ± j326	.304	1
M3	-.1	.5	.389 ± j392	1.047	2
M4	-.1	.75	.355 ± j477	1.039	0
M5	.3	.5	.497 ± j368	.808	3
M6	.5	.7	.496 ± j466	.736	1

Table 2

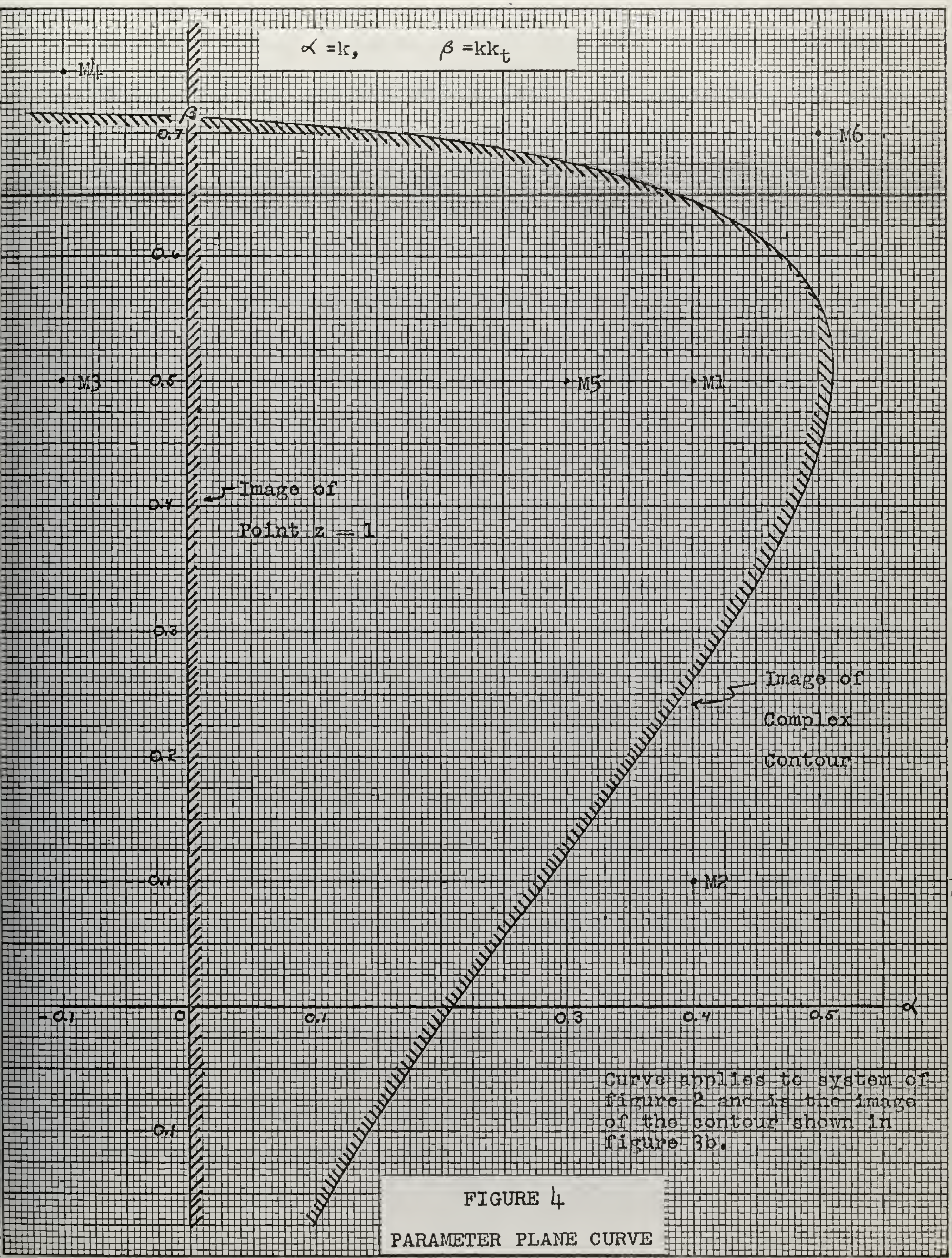


FIGURE 4

PARAMETER PLANE CURVE

achieved as follows. Recognize that the characteristic equation is of order three in z; further recognize that for $\alpha = \beta = 0$ the roots to the characteristic equation are identical to the open loop poles which are known to be within the contour. Hence all three roots to the characteristic equation lie within the designated contour for M-points located in the region which includes the parameter plane origin. Since this is the maximum number of roots which can be enclosed an M-point traveling across any boundary to this region implies roots leaving the chosen z-plane contour. With these observations shading can be effected without recourse to formulas). Once the parameter plane curves are drawn and shaded (as in figure 4) the compensation problem reduces to locating the M-point in the region which results in all roots lying within the specified z-plane contour. The preceding discussion indicates that this region is the one which includes the parameter plane origin. To verify this analysis six M-points in four different enclosed root regions have been indicated in figure 4, and the roots to the characteristic equation for each choice have been computed. Table 2 lists these roots and indicates how many fall within the specified z-plane contour. For each M-point theory and practice agree.

Let it be emphasized that this example illustrates parameter plane interpretation and shading procedures rather than servo-system design in the parameter plane. For the latter more contours (images) are drawn; shading may or may not be required and additional information to guide selection of an M-point may be displayed.

Mitrovic Curve Interpretation for Common Z-Plane Contours

Mitrovic methods (defined in chapter 2) are special cases of the parameter plane which warrant attention in linear (s-plane) theory because the Cauchy criterion regarding argument change for roots enclosed by s-plane

contours is directly interpretable as encirclements of the M-point. For the Mitrovic plane the tedious sub-division of the parameter plane into acceptable and inacceptable regions can be bypassed in favor of a simple inspection which determines whether the Cauchy criterion for enclosed roots is or is not complied with. The relation between constant zeta curves on the Mitrovic plane and the Cauchy criterion has been thoroughly explored in the literature. (6, 8). A compendium of these relations is simply this. When all roots of the characteristic equation are enclosed by the s-plane contour being mapped, argument f(s) where f(s) becomes the characteristic equation when equated to zero, increases monotonically. In turn this requires that the M-point be encircled in a predictable manner by the Mitrovic plane image of the s-plane contour. Were constant zeta contours meaningful in the z-plane, s-plane interpretation would be applicable to them; however, interesting z-plane contours are spirals of constant damping and circles of constant settling time. It can be shown that when all roots of the characteristic equation are within either of these contours argument f(z) increases monotonically as the contour is traversed in a positive direction. (Proof can be obtained by resolving f(z) into its factors and showing that each has a monotonically increasing argument if it is located within the contour). At this point complete analogy between s-plane and z-plane interpretation of Mitrovic curves fails. The monotonic increase in argument still sllows satisfaction of the Cauchy criterion to be determined from the manner in which image curves encircle the M-point in the Mitrovic (B_0-B_1) plane. However, a few simple examples suffice to show that this interpretation does not extend to generalized Mitrovic coefficients for z-plane contours of constant damping and constant settling time.

Recall that the characteristic equation and Mitrovic polynomial are defined as:

$$f(z) = F(z) + a_1 z + a_0$$

$$f_m(z) = F(z) + B_1(z)z + B_0(z) = 0$$

Subtraction of zero (the Mitrovic polynomial) from the first equation yields:

$$f(z) = (a_1 - B_1(z))z + (a_0 - B_0(z))$$

Consider the f(z) locus and its Mitrovic (B_1 - B_0) curve shown in figures 5a and 5b respectively.

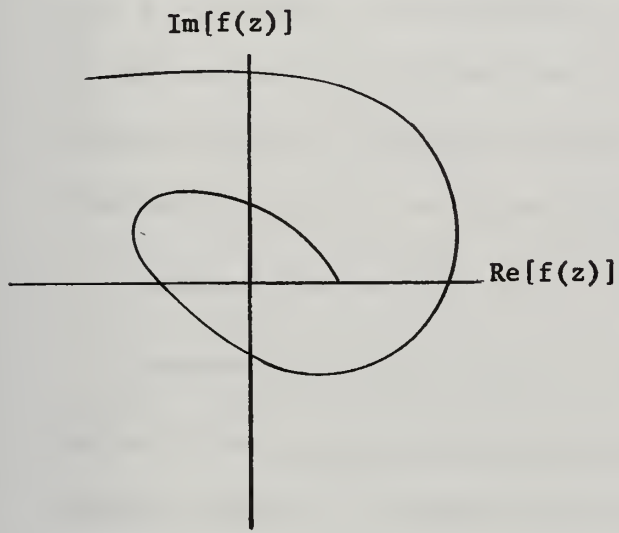


Figure 5a

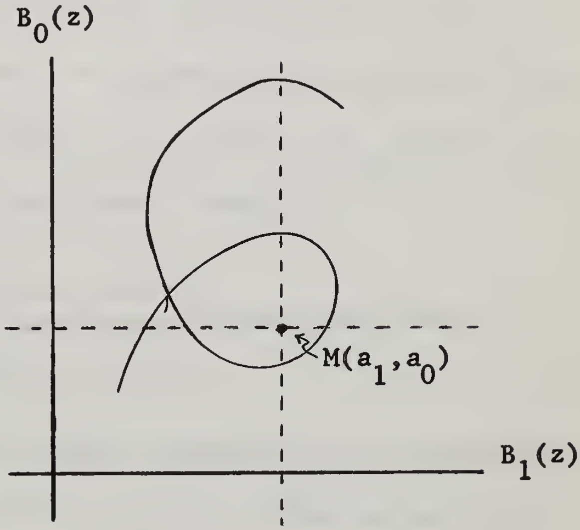


Figure 5b

If all zeros of f(z) are enclosed by the z-plane contour the f(z) locus always starts on the positive real axis and argument f(z) increases monotonically as the locus f(z) encircles the origin. The total change in argument f(z) is π radians times the number of roots of f(z) enclosed by the z-plane contour. Since argument f(z) is a continuous function (assuming no root of f(z) lies on the specified z-plane contour) the Cauchy criterion requires that

$$\text{argument } f(z_i) = k\pi \qquad k = 1, 2, 3, \ldots\ldots, n-1$$

(that is $f(z_i)$ be real) for n-1 values of z on the specified contour where f(z) is a polynomial of degree n. The n-1 values of z do not include the two points at which the specified contour intersects the real axis. At these two points f(z) must also be real but the interpretation which follows will not be valid. Take particular note that under these restrictions the imaginary part of z_i is not zero at these n-1 points. Hence

$$f(z) = [a_1 - B_1(z)]\, z_i + [a_0 - B_0(z_i)]$$

being real requires that $[a_1 - B_1(z_i)]z_i$ be real. This occurs only if $a_1 - B_1(z_i) = 0$; which in the Mitrovic plane requires that the B_0-B_1 curve lie on the vertical line through the M-point. Furthermore since argument f(z) is monotonically increasing the half lines $a_0 > B_0$ and $a_0 < B_0$ must be cut alternately with the half line $a_0 < B_0$ cut first. That is:

argument $f(z) = 2k\pi$	implies $a_0 > B_0$
argument $f(z) = (2k + 1)\pi$	implies $a_0 < B_0$

The half line $a_0 < B_0$ must be cut first because argument f(z) is monotonic and must equal π, a point on the half line $a_0 < B_0$, before it can equal 2π, the point at which it crosses the half line $a_0 > B_0$.

The points at which the specified z-plane contour intersects the real axis were excluded because previous arguments did not apply to them. Observe that if z_i is real

$$f(z_i) = [a_1 - B_1(z_i)]\, z_i + [a_0 - B_0(z_i)] \qquad \text{is real}$$

and there is no requirement that $[z_1 - B_1(z_i)] = 0$. In general this term is not equal to zero for real z and this fact can be used to determine the manner in which the B_0 -B_1 curve encircles the M-point. If z_r is the value at which the specified z-plane contour intersects the positive real axis then:

$B_1(z_r) > a_1$ implies counter clockwise encirclement

$B_1(z_r) < a_1$ implies clockwise encirclement

This follows directly from the requirement that the line $a_0 < B_0$ must be cut first. The value $B_1(z_r)$ can be computed by the relatively simple formula

$$B_1(z_r) = -\sum_{k=2}^{N} k\alpha_k z_r^{k-1}$$

Mitrovic $B_o - B_1$ curve interpretation for z-plane contours can be summarized.

1. If all roots of f(z) are located within a specified z-plane contour (settling circles or damping spirals) the $B_0(z)$ - $B_1(z)$ locus encircles the M-point in the Mitrovic plane.
2. The direction of encirclement is determined by:

   $B_1(z_r) > a_1$ implies counter clockwise encirclement

   $B_1(z_r) < a$ implies clockwise encirclement
3. The line $a_0 - B_0 = 0$ must be intersected (n-1) times plus one additional time each if the $B_0 - B_1$ curve starts or ends on this line.

In linear systems the next attempt was to extend Mitrovic's method to include any two coefficients of the characteristic equation. Desirable regions in the generalized plane were again found by inspection in the same manner as for the $B_o - B_1$ case. No such simple interpretation exists for z-transforms when generalized Mitrovic coordinates are used. Since zeta varies in all z-plane contours of interest the requirement that the B_i-B_j locus encircle the M-point (a_i-a_j) is no longer applicable. (i.e., the Cauchy criterion can no longer be interpreted in the B_i-B_j plane by inspection). A few illustrations will suffice to illustrate the point.

Illustration I: $f(z) = \left[a_2 - B_2(z)\right] z^2 + \left[a_0 - B_0(z)\right]$

The criterion "f(z) real" is met when z = j (without restrictions on the terms in brackets). Thus alternate cutting of the half lines $B_2 > a_2$ and $B_2 < a_2$ is not required and the M-point need not be encircled to meet the requirement that argument f(z) be monotonically increasing.

Illustration II: $f(z) = \left[\left\{ a_{k+1} - B_{k+1}(z) \right\} z + \left\{ a_k - B_k(z) \right\} \right] z^k$

The criterion "f(z) real" is satisfied if argument z^k equals minus argument $\left[\left\{ a_{k+1} - B_{k+1}(z) \right\} z + \left\{ a_k - B_k(z) \right\} \right]$. Again the argument requirement on f(z) cannot be determined by simple inspection of the $B_{k+1} - B_k$ curve.

Although it is probable that criteria for desired Mitrovic plane regions for generalized coefficients can be found, they will be complicated and it is recommended that parameter plane interpretation methods be used. And, without simple interpretation methods the Mitrovic plane loses its special attraction over the parameter plane and reverts to simply a special case of the latter.

CHAPTER IV

Preceding chapters endeavored to present and interpret the characteristic equation in a meaningful manner in the parameter plane. Parameter plane presentations need not be restricted to the characteristic equation; in a more general consideration the parameter plane is a space in which a system is described in terms of its two variable parameters. In analogy to characteristic equation presentation, system description in the parameter plane is the reverse of descriptions normally seen. Parameter plane presentation displays suitable values of α and β to meet a specified system performance criterion whereas conventional procedures first fix α and β and display system performance resulting from this choice. Significant advantages acrue when system criteria are described in the parameter plane.

1. The usual "trial and error" search of parameter settings to meet specifications is eliminated. This advantage is especially attractive when specifications are of the form "equal to or less (greater) than" and as such are not tractable mathematically. In these instances the locus satisfying equality usually divides the parameter plane into acceptable and unacceptable regions where parameter settings can be chosen by inspection to satisfy both this and other criteria.

2. Simultaneous consideration of all criteria and characteristic equation root information is accomplished by inspection. Tacitly assumed is that each criterion plots as a curve which divides the parameter plane into regions suitable or unsuitable for that criterion. Parameter settings

within the intersection of all suitable regions simultaneously meet all specifications.

Liabilities are also present in parameter plane presentations although they might more properly be classified as limitations.

1. System descriptions are limited to description in two parameters (even though additional parameters may be available for adjustment).
2. Specified criteria may be difficult or impossible to determine as functions of adjustable system parameters. A maximum overshoot specification typifies such criteria. Theoretically a locus of α-β plane points which have the specified maximum overshoot exists but in practice the locus is virtually impossible to determine. Time-integral criteria and magnitude of frequency response at resonance also seem to fall within this class.

These limitations not withstanding, two very important performance specifications, steady state error and bandwidth, can be displayed on the parameter plane. They complement parameter plane information already available from characteristic equation roots by relating system performance to the zeros of the closed loop transfer function as well as to the poles. (The formal consideration of closed loop zeros is almost universally neglected by other transform analysis techniques).

Steady State Error Curves in the Parameter Plane

The steady state error or inability of a servo-system to follow an aperiodic input disturbance is an important criterion and frequent system specification. For linear systems or linear sampled data systems steady state error is usually obtained by application of the final value theorem

to the LaPlace or z-transform of the error function. This theorem also allows the requisite steady state error curves to be computed for parameter plane display. Derivation of the equation governing these curves in the parameter plane will be for the z-transformation with a subsequent paragraph indicating modifications required to extend applicability to the LaPlace transformation and continuous systems.

Assume that the open loop transfer function of a unity feedback system (or a system which has been reduced to unity feedback form) is

$$\frac{N_o(z)}{(z-1)^{\ell} D_o(z)}$$

where $D_o(z)$ contains no factors of (z-1) nor factors in common with $N_o(z)$. Further assume that both $N_o(z)$ and $D_o(z)$ contain no terms above products in the parameters α and β. These assumptions are illustrated in figure 6 and equations (33).

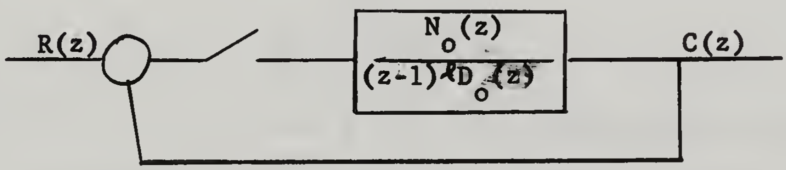


Figure 6

$$N_o(z) = \sum_{k=0}^{N} \left(\alpha p_k^o + \beta q_k^o + \alpha\beta g_k^o + h_k^o\right) z^k \qquad (33)$$

$$D_o(z) = \sum_{k=0}^{N} \left(\alpha b_k + \beta c_k + \alpha\beta d_k + e_k\right) z^k$$

Typical steady state error specifications require aperiodic inputs of steps, ramps, accelerations, whose z-transforms have the form

$$R(z) = \frac{R^*(z)}{(z-1)^m}$$

where $R^*(z)$ contains no factor of (z-1) or $(z-1)^{-1}$

m = 1 implies step inputs, m = 2 implies ramp inputs and

m = 3 implies acceleration inputs.

The z-transformed error signal to which the final value theorem will be applied is obtained by block diagram manipulation.

$$E(z) = (z-1)^{m-\ell} \frac{D_o(z) R^*(z)}{N_o(z) + (z-1)^m D_o(z)}$$

The final value theorem then gives the steady state error as

$$E_{s.s.} = \lim_{z \to 1} \left[\frac{(z-1)^{\ell+1-m}}{z} \cdot \frac{D_o(z) R^*(z)}{N_o(z) + (z-1)^m D_o(z)} \right] \tag{34}$$

Consideration of equation (34) is divided into three cases.

1. $\ell + 1 - m \neq 0$ In this event the factor of (z-1) or $(z-1)^{-1}$ present approaches zero or infinity as the limit is passed and error specifications have no meaning for the input applied.

2. $\ell + 1 - m = 0$ and $\ell = 0$ This denotes a type zero system in which the steady state error for a step input is

$$E_{s.s.} = \frac{D_o(1) R^*(1)}{N_o(1) + D_o(1)} \tag{35}$$

3. $\ell + 1 - m = 0$ and $\ell \neq 0$ This denotes type I systems and above in which a factor approaching zero as a limit multiplies $D_o(1)$ in the denominator. The steady state error expression appropriate here is

$$E_{s.s.} = \frac{D_o(1) R^*(1)}{N_o(1)} \tag{36}$$

Although their present statement does not so indicate equations (35) and (36) are the relations which permit steady state error specifications to be drawn on the parameter plane. Statements adapted to parameter plane use can be obtained by evaluating $D_o(1)$ and $N_o(1)$ in equations (33) and defining a more concise notation for the resulting sums. That is:

$$N_o(1) = \sum_{K=0}^{N} \left(\alpha p_K^o + \beta q_K^o + \alpha\beta g_K^o + h_K^o \right)$$

$$D_o(1) = \sum_{K=0}^{N} \left(\alpha b_K^o + \beta c_K^o + \alpha\beta d_K^o + e_K^o \right)$$

Now define:

$$E_B = \sum_{K=0}^{N} b_K^o \qquad E_P = \sum_{K=0}^{N} p_K^o$$

$$E_C = \sum_{K=0}^{N} c_K^o \qquad E_Q = \sum_{K=0}^{N} q_K^o$$

$$E_D = \sum_{K=0}^{N} d_K^o \qquad E_G = \sum_{K=0}^{N} g_K^o \tag{37}$$

$$E_E = \sum_{K=0}^{N} e_K^o \qquad E_H = \sum_{K=0}^{N} h_K^o$$

With these definitions $N_o(1)$ can be expressed as

$$N_o(1) = \alpha E_P + \beta E_Q + \alpha\beta E_G + E_H$$

$$D_o(1) = \alpha E_B + \beta E_C + \alpha\beta E_D + E_E \tag{38}$$

where all E's are constants. Useable expressions for steady state error curves on the α-β plane are rendered by substitution of expression (38) into equations (35) and (36). For substitution into (36)

$$E_{s.s.} = \frac{(\alpha E_B + \beta E_C + \alpha\beta E_D + E_E) R^*(1)}{(\alpha E_P + \beta E_Q + \alpha\beta E_G + E_H)}$$

In this expression all E's are constants as is R*(1). When $E_{s.s.}$ is equated to the steady state error specification the equation describes the locus of all α-β plane points exactly meeting the requirement. The locus explicitly expressed for β as a function of α is

$$\beta = - \frac{\alpha \left[E_B R^*(1) - E_P E_{s.s.} \right] + \left[E_E R^*(1) - E_H E_{s.s.} \right]}{\alpha \left[E_D R^*(1) - E_G E_{s.s.} \right] + \left[E_C R^*(1) - E_Q E_{s.s.} \right]} \tag{39}$$

For type I systems and above equation (39) defines the locus of all points (α, β) having a steady state error $E_{s.s.}$ to an input whose z-transform

is $\frac{R(z)}{(Z-1)^{\ell}}$.

For type zero systems an analogous formula which describes steady state error loci for an input step of magnitude R can be obtained by initially substituting expressions (38) into equation (35). The final result is

$$\beta = -\frac{\alpha\left[E_B(R-E_{s.s.})-E_P E_{s.s.}\right]+\left[E_E(R-E_{s.s.})-E_H E_{s.s.}\right]}{\alpha\left[E_D(R-E_{s.s.})-E_G E_{s.s.}\right]+\left[E_C(R-E_{s.s.})-E_\phi E_{s.s.}\right]} \tag{40}$$

Observe that the steady state error loci given by equations (39) and (40) describe hyperbolas in the parameter plane for the $\alpha\beta$ product case. For the linear case $E_D \equiv E_G \equiv 0$ and the steady state error loci become straight lines in the parameter plane.

Steady State Error Loci for LaPlace Transformed Variables

If anything steady state error loci are easier to obtain from error signals described by their LaPlace transform as opposed to signals described by their z-transform. As for the sampled data case assume a known open loop transfer function with numerator and denominator linear in α and β but not necessarily linear with respect to each other. That is:

$$N_o(s) = \sum_{K=0}^{N}\left(\alpha p_K^o + \beta q_K^o + \alpha\beta g_K^o + h_K^o\right) S^K \tag{41}$$

$$D_o(s) = \sum_{K=0}^{N}\left(\alpha b_K^o + \beta c_K^o + \alpha\beta d_K^o + e_K^o\right) S^K$$

where $D_o(s)$ does not contain s as a factor and the complete denominator is

$$s^{\ell}\, D_o(s)$$

Again consider only inputs of the step, ramp, acceleration type.

Input = Rt^{m-1} $\qquad\qquad$ $\mathcal{L}\{\text{input}\} = \frac{(M-1)!\,R}{S^M}$

Under these assumptions and definitions the LaPlace transform of the error signal is:

$$E(s) = \frac{S^{\ell} D_o(S)}{N_o(S) + S^{\ell} D_o(S)} \cdot \frac{(M-1)!\,R}{S^M}$$

For stable systems the final value theorem can be applied to exhibit steady state error as:

$$E_{s.s.} = \lim_{S\to 0} \left[S\,E(S)\right] = \lim_{S\to 0}\left[S^{\ell+1-m}\,\frac{(m-1)!\,R\,D_o(S)}{N_o(S)+S^{\ell}D_o(S)}\right] \tag{42}$$

On passing the limit results can again be divided into three cases.

1. $\ell + 1 - m \neq 0$ In this case steady state error is either zero or infinite and no finite adjustment of the parameters can alter this result.

2. $\ell + 1 - m = 0$, $\ell \neq 0$ These requirements describe a system of type I or higher. The parameter plane locus of points having a steady state error of $E_{s.s.}$ as obtained from equations (41) and (42) is:

$$\beta = -\frac{\alpha\left[b_o^{\circ}(m-1)!\,R - p_o^{\circ}E_{s.s.}\right] + \left[e_o^{\circ}(m-1)!\,R - h_o^{\circ}E_{s.s.}\right]}{\alpha\left[d_o^{\circ}(m-1)!\,R - g_o^{\circ}E_{s.s.}\right] + \left[c_o^{\circ}(m-1)!\,R - q_o^{\circ}E_{s.s.}\right]} \tag{43}$$

3. $\ell + 1 - m = 0$, $\ell = 0$ This requirement is indicative of type zero systems which are characterized by having steady state errors to step inputs. The parameter plane loci for a steady state error $E_{s.s.}$ to an input step of magnitude R are governed by the equation:

$$\beta = -\frac{\alpha\left[b_o^{\circ}(R-E_{s.s.}) - p_o^{\circ}E_{s.s.}\right] + \left[e_o^{\circ}(R-E_{s.s.}) - h_o^{\circ}E_{s.s.}\right]}{\alpha\left[d_o^{\circ}(R-E_{s.s.}) - g_o^{\circ}E_{s.s.}\right] + \left[c_o^{\circ}(R-E_{s.s.}) - q_o^{\circ}E_{s.s.}\right]} \tag{44}$$

Observe that application of the final value theorem makes only those terms in the numerator and denominator not multiplied by s (after s factors in numerator and denominator are removed) significant in determination of steady state error. If parameters α and β do not appear in these terms steady state error is not influenced by their adjustment regardless of where they appear elsewhere. Furthermore, only the manner in which α and β

[illegible]

appear in these terms affects the shape of the steady state error curves. If $\alpha\beta$ products appear in these terms the steady state error loci will be hyperbolas; if α and β appear linearly the loci will be straight lines.

Bandwidth Curves in the Parameter Plane

> The bandwidth is significant because it indicates rise time or speed of response, it measures in part the ability of the system to reproduce the input signal, and it approximately describes the filtering characteristics of the system.[1]

Bandwidth specifications can be presented on the parameter plane with no restrictions above those inherent to explicit solution of α and β When so displayed, bandwidth can be considered in conjunction with other specifications presentable on the parameter plane and satisfaction of all is virtually by inspection, if such satisfaction is possible. Before embarking on the derivation of formulas which display bandwidth criteria on the parameter plane some preliminaries deserve attention. Among these is agreement on a definition for bandwidth.

> Excite a servo-system by a sinusoidal test signal and form the ratio of peak fundamental component output to peak input. (If the system is linear the output consists wholly of a fundamental component). The lowest frequency at which this ratio falls below a designated value A is defined as the bandwidth of the servo-system.

The most widely accepted value for A is 0.707 (-3 db) although unity and other values have been used. A readily accessible expression from which the ratio of peak fundamental output to peak input can be computed is the closed loop transfer function. If Fourier transforms are employed the magnitude of the closed loop transfer function represents this ratio. The identical ratio is obtainable from the magnitude of the closed loop transfer function

[1] Truxal, J. G. Control Systems Synthesis. McGraw-Hill, 1955: 77.

expressed in terms of the complex LaPlace variable s if the substitution $s = jw$ is effected. For sampled signals a similar ratio is obtained by examining the closed loop z-transfer function for values of z on the unit circle.

Parameter plane methods are not primarily concerned with a determination of this ratio; rather the values of system variables (locus of α-β plane points) which meet bandwidth specifications (cause the output-input ratio to be of proper magnitude at the specified frequency) are sought. Functionally the parameter plane bandwidth problem is expressed:

$$A = \left| \frac{C(\alpha, \beta, \omega_{BW})}{R(\omega_{BW})} \right| \tag{45}$$

where by adjustment of system parameters α and β the magnitude of the closed loop transfer function is forced to have a magnitude A at the bandwidth frequency $w_{b.w.}$. Equation (45) while sufficient to define bandwidth loci on the parameter plane is not conveniently utilized for their construction. The remainder of the section is devoted to remedy of this problem.

In line with parameter plane assumptions the closed loop system transfer function is required to have the form

$$\frac{C(z)}{R(z)} = \frac{\sum_{K=0}^{N} (\alpha p_K + \beta q_K + \alpha\beta g_K + h_K) z^K}{\sum_{K=0}^{N} (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K) z^K} \tag{46}$$

In the course of the derivation it proves convenient to separate both the numerator and the denominator into their real and imaginary components. The Chebyshev functions introduced in equations (5) through (9) provide a satisfactory vehicle for this purpose. For reference, z defined in terms of the Chebyshev functions is

$$z = \omega_{N_z} \left[\zeta_z + j\sqrt{1-\zeta_z^2} \right]$$
$$z^k = \omega_{N_z}^k \left[T_K(\zeta_z) + j\sqrt{1-\zeta_z^2}\, U_K(\zeta_z) \right]$$

To promote brevity the real and imaginary parts of each sum are defined as:

$$R_P(\zeta_z, \omega_{N_z}) = \sum_{K=0}^{N} p_K \omega_{N_z}^K T_K(\zeta_z)$$
$$I_Q(\zeta_z, \omega_{N_z}) = \sum_{K=0}^{N} q_K \omega_{N_z}^K \sqrt{1-\zeta_z^2}\, U_K(\zeta_z)$$
$$R_D(\zeta_z, \omega_{N_z}) = \sum_{K=0}^{N} d_K \omega_{N_z} T_K(\zeta_z) \tag{47}$$
$$I_E(\zeta_z, \omega_{N_z}) = \sum_{K=0}^{N} e_K \omega_{N_z} \sqrt{1-\zeta^2}\, U_K(\zeta_z)$$

The remaining twelve quantities in this set are similarly defined. It is noteworthy that by this definition all quantities R and I are real. On replacing the sums in equation (46) by their defined equivalents in equation (47) one obtains:

$$\frac{C(z)}{R(z)} = \frac{\alpha(R_P + jI_P) + \beta(R_Q + jI_Q) + \alpha\beta(R_G + jI_G) + (R_H + jI_H)}{\alpha(R_B + jI_B) + \beta(R_C + jI_C) + \alpha\beta(R_D + jI_D) + (R_E + jI_E)} \tag{48}$$

Bandwidth specification that the ratio of sampled output to sampled input have some magnitude A at a bandwidth frequency of $w_{b.w.}$ equates to the requirement that $C(z_{b.w.})$ over $R(z_{b.w.})$ have magnitude A where $z_{b.w.}$ is the value of z on the unit circle which images the bandwidth frequency $s = jw_{b.w.}$ in the s-plane. When these bandwidth conditions are introduced the quantities R and I become constants and equation (48) describes the locus of all points (α, β) which satisfy system bandwidth specifications. That is:

$$A = \left| \frac{C(z_{b.w.})}{R(z_{b.w.})} \right|$$

$$A^2 = \frac{(\alpha R_P + \beta R_Q + \alpha\beta R_G + R_H)^2 + (\alpha I_P + \beta I_Q + \alpha\beta I_G + I_H)^2}{(\alpha R_B + \beta R_C + \alpha\beta R_D + R_E)^2 + (\alpha I_B + \beta I_C + \alpha\beta I_D + I_E)^2}$$

where A is the magnitude of attenuation specified at the bandwidth frequency.

All R's and I's are constants.

Performing indicated multiplications allows this equation to be rendered as:

$$A^2 = \frac{\alpha^2\beta^2(R_G^2+I_G^2) + 2\alpha^2\beta(R_PR_G+I_PI_G) + 2\alpha\beta^2(R_QR_G+I_QI_G) +}{\alpha^2\beta^2(R_D^2+I_D^2) + 2\alpha^2\beta(R_BR_D+I_BI_D) + 2\alpha\beta^2(R_CR_D+I_CI_D) +} \tag{49}$$

$$\frac{\alpha^2(R_P^2+I_P^2) + 2\alpha\beta(R_PR_Q+R_GR_H+I_PI_Q+I_GI_H) + \beta^2(R_Q^2+I_Q^2) +}{\alpha^2(R_B^2+I_B^2) + 2\alpha\beta(R_BR_C+R_DR_E+I_BI_C+I_DI_E) + \beta^2(R_C^2+I_C^2) +}$$

$$\frac{2\alpha(R_PR_H+I_PI_H) + 2\beta(R_QR_H+I_QI_H) + (R_H^2+I_H^2)}{2\alpha(R_BR_E+I_BI_E) + 2\beta(R_CR_E+I_CI_E) + (R_E^2+I_E^2)}$$

Equation (49) describes the locus of parameter plane points which satisfy a given bandwidth specification demanding an attenuation A at bandwidth frequency $w_{b.w.}$. Its appearance is forbidding but after recognition of certain features the equation itself is not formidable. Foremost is that α and β individually appear to no power higher than two. Thus to find a point on this bandwidth locus for a given alpha value involves no more than solution for β by the quadratic formula. (Computer plotting of such a curve is trivial). It is also significant that equation (49) describes a conic section whenever the numerator and denominator coefficients are linear in α and β in which case $R_G = I_G = R_D = I_D \equiv 0$.

To demonstrate the first of these observations explicit solution for β will be effected under the assumption that α is known (assigned). For convenience first define:

$$X_1(\alpha) = \alpha^2(R_G^2 + I_G^2) + 2\alpha(R_QR_G + I_QI_G) + (R_Q^2 + I_Q^2)$$

$$X_2(\alpha) = \alpha^2(R_D^2 + I_D^2) + 2\alpha(R_CR_D + I_CI_D) + (R_C^2 + I_C^2)$$

$$Y_1(\alpha) = 2\alpha^2(R_PR_G + I_PI_G) + 2\alpha(R_PR_Q + R_GR_H + I_PI_Q + I_GI_H) + 2(R_QR_H + I_QI_H) \tag{50}$$

$$Y_2(\alpha) = 2\alpha^2(R_BR_D + I_BI_D) + 2\alpha(R_BR_C + R_DR_E + I_BI_C + I_DI_E) + 2(R_CR_E + I_CI_E)$$

$$W_1(\alpha) = \alpha^2(R_P^2 + I_P^2) + 2\alpha(R_PR_H + I_PI_H) + (R_H^2 + I_H^2)$$

$$W_2(\alpha) = \alpha^2(R_B^2 + I_B^2) + 2\alpha(R_BR_E + I_BI_E) + R_E^2 + I_E^2)$$

Explicit solution of β from equation (49) utilizing the notation defined in (50) is:

$$\beta = \frac{(A^2Y_2 - Y_1) \pm \sqrt{(A^2Y_2 - Y_1)^2 - 4(A^2X_2 - X_1)(A^2W_2 - W_1)}}{2(A^2X_2 - X_1)} \tag{51}$$

Equation (51) permits determination of the β coordinate of parameter plane points meeting a bandwidth specification given the value of the α coordinate. A negative value under the radical sign infers that no point β will satisfy the bandwidth specification for the value of α assigned. When bandwidth is specified at a frequency of $w = {}^{w}s/_{2}$ the bandwidth loci in the parameter plane degenerate from conic sections into straight lines (for the linear coefficient case). Whether interpretation of these degenerate curves is consistent with interpretation of bandwidth curves specified at other frequencies has not been established; hence, the reader is cautioned to avoid specifying bandwidth at the frequency of ${}^{w}s/_{2}$.

<u>Bandwidth Curves for S-Plane Transfer Functions</u>

The preceding bandwidth theory is applicable to closed loop transfer functions described in terms of the complex variable s. In fact only the definitions of the R's and I's need be altered to make z-plane bandwidth equations valid for s-plane use. For the sake of completeness the derivation below duplicates some of the work in the preceding section. Assume a

closed loop transfer function:

$$\frac{C(S)}{R(S)} = \frac{\sum_{K=0}^{N} (\alpha p_K + \beta q_K + \alpha\beta g_K + h_K) S^k}{\sum_{K=0} (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K) S^k}$$

$$\frac{C(S)}{R(S)} = \frac{\sum_{K=0}^{N} (\alpha p_K + \beta q_K + \alpha\beta g_K + h_K) \omega_N^k \left[T_K(-\zeta) + j\sqrt{1-\zeta^2}\, U_K(-\zeta)\right]}{\sum_{K=0} (\alpha b_K + \beta c_K + \alpha\beta d_K + e_K) \omega_N^k \left[T_K(-\zeta) + j\sqrt{1-\zeta^2}\, U_K(-\zeta)\right]}$$

Since bandwidth is always specified at some real frequency only zeta values of zero are relevant. For = 0 the $T_k(0)$ and $U_k(0)$ are expressible as:

$$T_k(0) = \cos \frac{k\pi}{2} \qquad\qquad U_k(0) = \sin \frac{k\pi}{2}$$

This observation allows definitions of the R's and I's in equations (47) to be altered for s-plane use.

$$\begin{aligned}
R_P(w_n) &= \sum_{K=0}^{N} (-1)^k p_{2k}\, \omega_N^{2k} & I_P(w_n) &= \sum_{K=0}^{N} (-1)^k p_{2k+1}\, \omega_N^{2k+1} \\
R_G(w_n) &= \sum_{K=0}^{N} (-1)^k g_{2k}\, \omega_N^{2k} & I_G(w_n) &= \sum_{K=0}^{N} (-1)^k g_{2K+1}\, \omega_N^{2k+1} \quad (52) \\
R_E(w_n) &= \sum_{K=0}^{N} (-1)^k e_{2K}\, \omega_N^{2k} & I_E(w_n) &= \sum_{K=0}^{N} (-1)^k e_{2K+1}\, \omega_N^{2k+1}
\end{aligned}$$

Omitted R's and I's are analogously defined.

S-plane bandwidth specifications require the magnitude of output to input to have suffered an attenuation A at the bandwidth frequency $w_{b.w.}$.

Functionally:

$$A = \frac{|C(j\omega_{bw})|}{|R(j\omega_{bw})|}$$

Evaluate equations (52) at $w_n = w_{b.w.}$ and substitute the result into the equation above to obtain:

$$A = \frac{|\alpha (R_P + jI_P) + \beta (R_Q + jI_Q) + \alpha\beta (R_G + jI_G) + (R_H + jI_H)|}{|\alpha (R_B + jI_B) + \beta (R_C + jI_C) + \alpha\beta (R_D + jI_D) + (R_E + jI_E)|} \qquad (53)$$

From this point all equations (49) and subsequent for z-plane bandwidth determination apply identically for s-plane bandwidth determination.

Periodicity of Bandwidth Problem

A periodic frequency spectrum is characteristic of any signal described by z-transforms. Frequently this periodicity characteristic is not intrinsic to the signal being described but is artifically introduced by a fictitious sampler to facilitate mathematical analysis. The commonplace situation for sampled data systems is illustrated in figure (7).

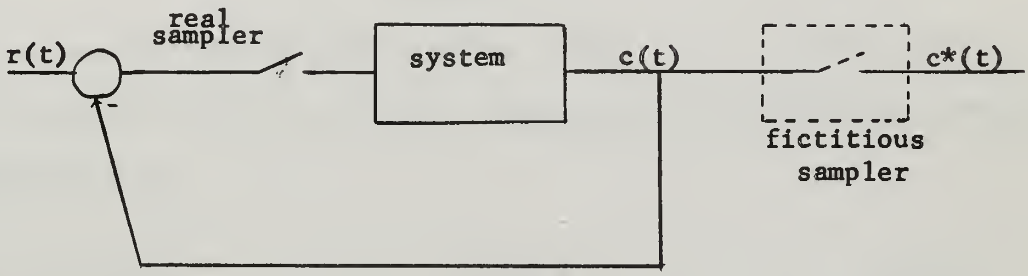


Figure 7

This frequency spectrum periodicity present in z-transfer functions radically alters frequency response interpretation and invalidates or restricts prevalent frequency response criteria when applied to pulsed transfer functions. Among criteria so altered is bandwidth; the 3 db attenuation usually specified for servo-system bandwidth may never occur when testing the pulsed transfer function yet the system itself may be perfectly satisfactory. This dilemma can profitably be bypassed by return to system description in the s-plane even though such description involves transcendental functions. Here bandwidth still has meaning and attention need not be restricted to sampled signals, a significant advantage when examining the system output.

Sampling makes the previous definition of bandwidth somewhat inadequate (as will subsequently be shown) and agreement on a suitable definition of bandwidth is incumbent. Consider the error sampled system in

figure (8).

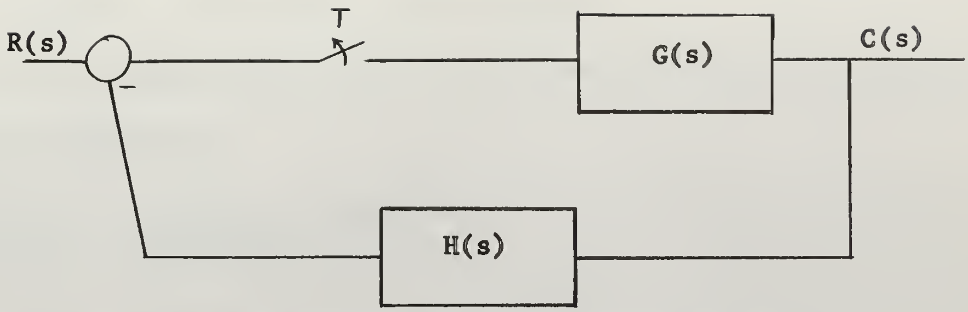


Figure 8

Denote by superscript X*(s) the LaPlace transform of a sampled function whose continuous transform is X(s). Reasonable alternatives for definition of bandwidth are:

$$A = \frac{|C^*(s)|}{|R^*(s)|}\bigg|_{s=j\omega_{bw}} \qquad A = \frac{|C(s)|}{|R(s)|}\bigg|_{s=j\omega_{bw}} \qquad A = \frac{|C(s)|}{|R^*(s)|}\bigg|_{s=j\omega_{bw}}$$

The first of these alternatives is identical to results obtained using z-transfer functions and its inadequacy has already been discussed. The second of these alternatives allows a different bandwidth for inputs $R_1(s)$ and $R_2(s)$ having the same value at the sampling instants. For the error sampled system shown the output is unaffected by the value of the input at other than the sampling instants and one suspects that any definition of bandwidth should also have this property. Therefore bandwidth appears most suitably defined when stated as:

$$A = \frac{|C(s)|}{|R^*(s)|}\bigg|_{s=j\omega_{bw}}$$

Even this statement needs minor modification to be intuitively satisfying. The example in the following section will demonstrate a more satisfactory and inherently pleasing definition for bandwidth is:

$$A = \frac{1}{T}\frac{|C(s)|}{|R^*(s)|}\bigg|_{s=j\omega_{bw}} \tag{54}$$

It is acknowledged that this definition for bandwidth is subject to debate when a system is not input or error sampled.

Block diagram manipulation of the system in figure (8) shows the input to be related by

$$C(s) = \frac{G(S)}{1+GH^*(S)} R^*(S)$$

This relation allows expression of system bandwidth as:

$$A = \frac{1}{T}\left|\frac{G(S)}{1+ GH^*(S)}\right| \tag{55}$$

A proper approach toward rendering this bandwidth expression fit for parameter plane use might be to examine its individual terms. G(s) is the system's LaPlace transfer function which, with the possible exception of a hold element, is presumed linear. As such G(s) must be a ratio of two polynomials in "s" and, if a hold element is included, e^{sT}. GH*(s) is the LaPlace transform of a sampled function and it can be represented as an infinite series.

$$X^*(s) = \sum_{N=0}^{\infty} X(NT)\, e^{-NTs}$$

where x(t) is the impulse response of component X(s). Whenever G(s) and H(s) are linear except for hold circuits, as is assumed, the series for GH*(s) has a closed form representation which is a rational function of e^{sT}. Significantly this closed form is readily obtainable from z-transform tables.

$$GH^*(s) \equiv GH(z)\Big|_{z=e^{sT}} \tag{56}$$

Also noteworthy is that specification of bandwidth at a frequency of $s = jw_{B.W.} = jw_s/2$ forces GH*(s) to be real. (i.e. Specify an input-output

attenuation of magnitude A at $w_s/2$ in preference to specification of an attenuation A' at some other frequency). Such specification reduces the problem of separating the real and imaginary parts of the function

$$\frac{G(S)}{1 + GH^*(S)}$$

preparatory to obtaining its magnitude, to a problem of separating the real and imaginary parts of G(jw) which is a rational function of w.

Still requisite is an equation which allows ready determination of parameter plane points satisfying the bandwidth specification of equation (55). Assume that the linear or product appearance of α and β in the closed loop transfer function continues when bandwidth is as specified in equation (55). To elaborate, grant that systems having closed loop transfer functions shown in equation (46) will retain this same type α-β dependence under equation (55).

$$\frac{C(S)}{R^*(S)} = \frac{\alpha f_P(S) + \beta f_Q(S) + \alpha\beta f_G(S) + f_H(S)}{\alpha f_B(S) + \beta f_C(S) + \alpha\beta f_D(S) + f_E(S)}$$

Evaluate the functions f at the specified bandwidth frequency $s = jw_{B.W.}$ and define:

$$R_P = \text{Real Part } [f_P(jw_{B.W.})]$$

$$I_P = \text{Imaginary Part } [f_P(jw_{B.W.})] \qquad (57)$$

$$R_B = \text{Real Part } [f_B(jw_{B.W.})]$$

$$I_B = \text{Imaginary Part } [f_B(jw_{B.W.})]$$

All R's and I's omitted above are similarly defined. Also define:

$$A' = A^{\cdot}T \qquad (57a)$$

Parameter plane bandwidth loci for bandwidth as defined in (55) can be obtained using definitions (57) and equations (49) and subsequent which were originally developed for z-plane operations.

Illustration of Auxiliary Parameter Plane Curves

The generalization that examples always serve to clarify theory is presumptuous especially in this instance when bandwidth and steady state error loci are already known to be conic sections and straight lines respectively. Nevertheless an example is presented and surprisingly the generalization holds. In fact a great deal can be learned from this example if digression to include several closed loop frequency response curves is allowed. Consider the third order sampled data system illustrated in figure (2) and direct attention to its closed loop frequency curves $\left|\frac{C(j\omega)}{R^*(j\omega)}\right|$ for representative parameter settings of

a) $k = 0.5$, $k_t = 1.0$ figure (8a)

b) $k = 0.1$ $k_t = 1.0$ figure (8b)

These curves are drawn in figure (8). Immediately obvious and disturbing is the fact that the closed loop d.c. ($w=0$) response of this feedback system has a magnitude differing from unity. One intuitively expects eventual equality of output and input when a type I system is disturbed by a d.c. signal; an expectation which equates to a frequency response of unit magnitude for $w=0$. The cause for fallacy of intuition and a definition modification which results in more pleasing (familiar) frequency response magnitude curves are obtainable from examination of the relative magnitudes of R(s) and R*(s). To expand, the magnitude of the frequency response as defined for continuous systems and also the definition on which intuition is based is

$$M = \left|\frac{C(j\omega)}{R(j\omega)}\right|$$

whereas a wiser choice for sampled data systems seems to be

$$M' = \left|\frac{C(j\omega)}{R^*(j\omega)}\right|$$

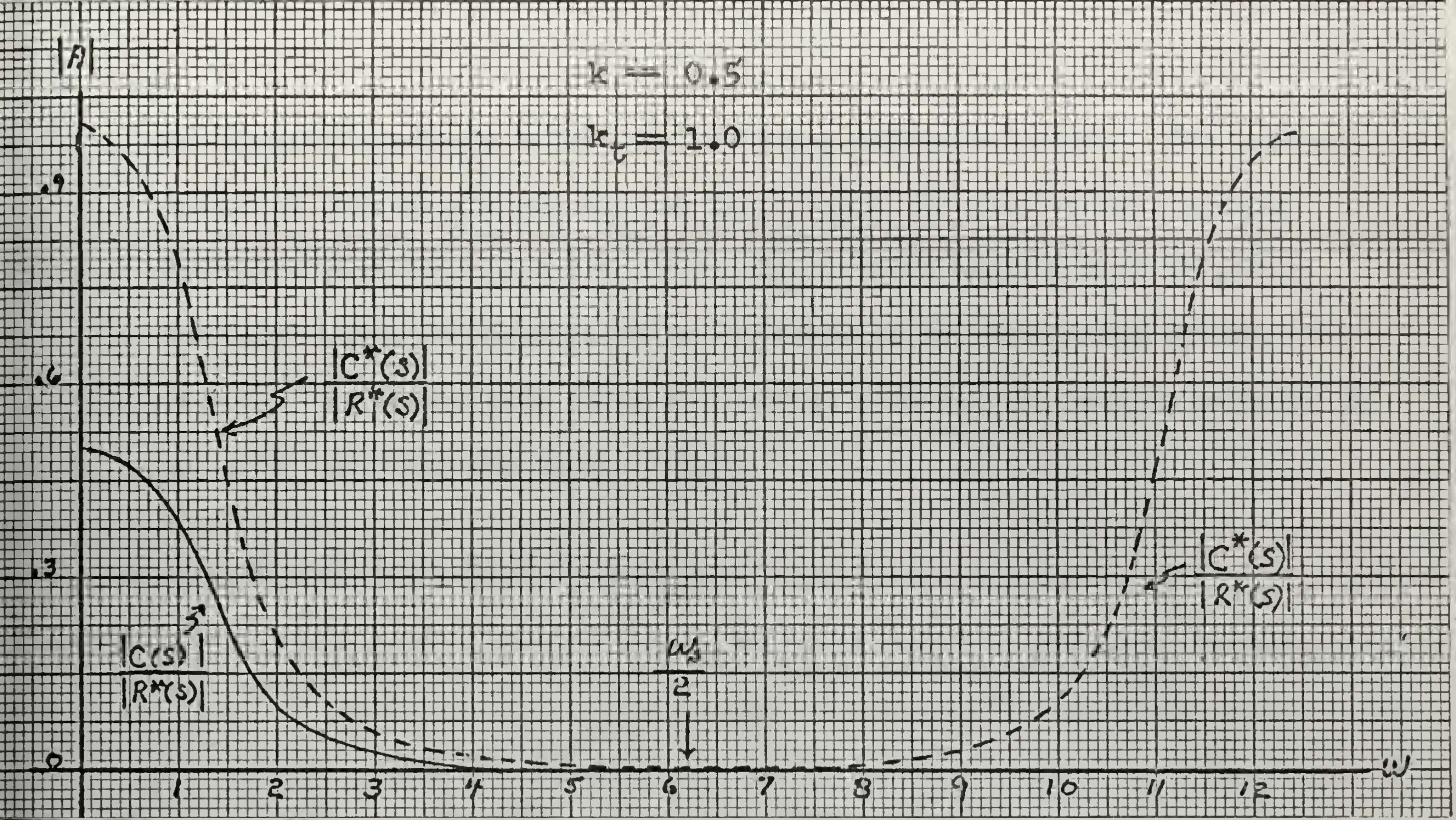


FIGURE 8a

FREQUENCY RESPONSE CURVES

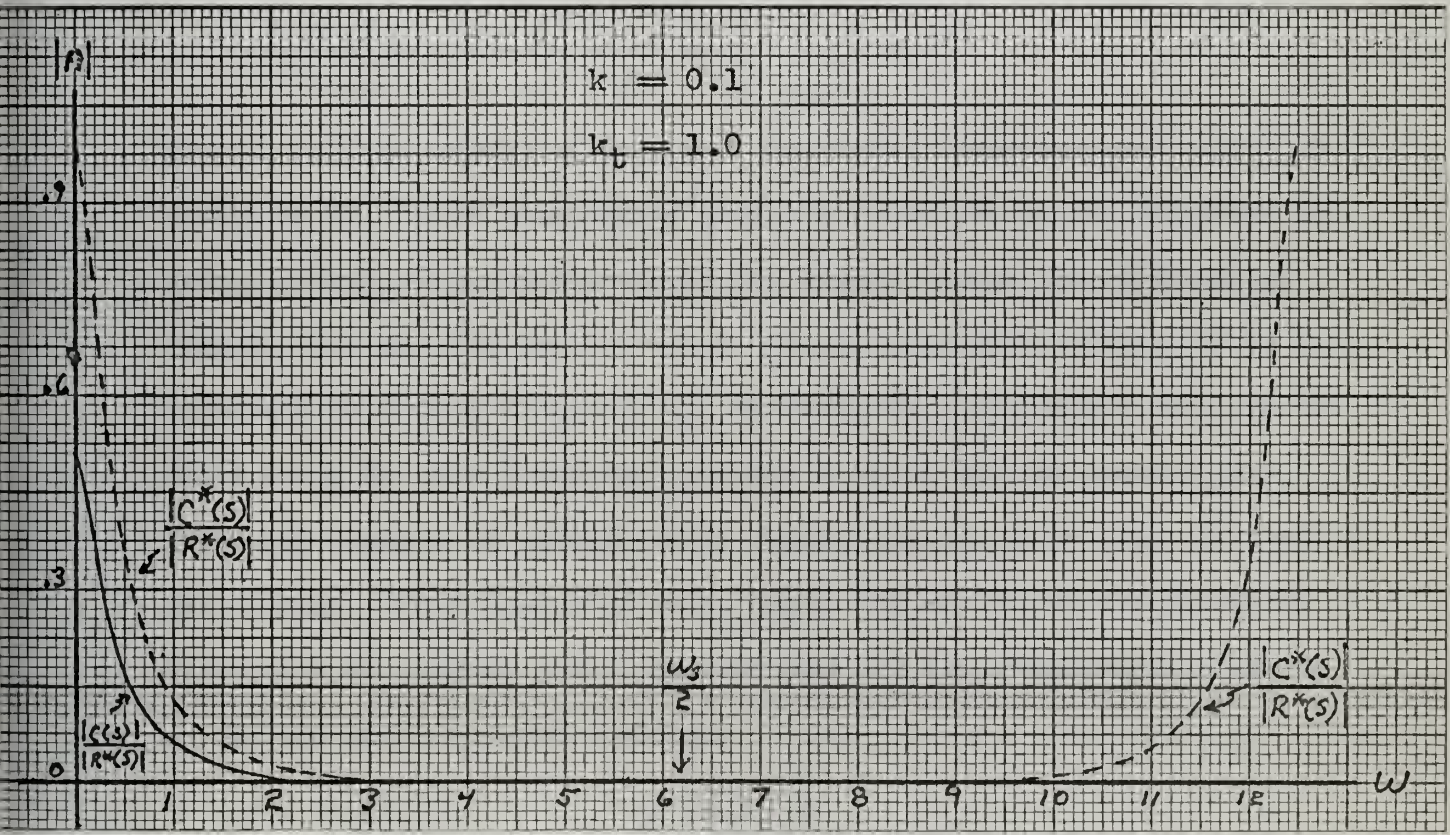


FIGURE 8b

FREQUENCY RESPONSE CURVES

Now suppose magnitude R(jw) does not equal magnitude R*(jw). Then correct intuitive reasoning precludes the same result for sampled data systems. However a scale factor may be available to force sampled data frequency response curves to résemble their counterparts for continuous systems more closely. To investigate these conjectures consider the general harmonic input $e^{jw_0 t}$ and form the ratio $\frac{R(s)}{R^*(s)}$

$$\mathcal{L}\{r(t)\} = R(s) = \frac{1}{s - j\omega_0}$$

$$\mathcal{L}\{r^*(t)\} = R^*(s) = \frac{e^{sT}}{e^{sT} - e^{j\omega_0 T}} = \frac{1}{1 - e^{-(s-j\omega_0)T}}$$

$$\frac{R(s)}{R^*(s)} = \frac{1 - e^{-(s-j\omega_0)T}}{s - j\omega_0}$$

The magnitude of this ratio is of interest at the frequency of the forcing function only.

$$\left|\frac{R(j\omega_0)}{R^*(j\omega_0)}\right| = \left|\frac{1 - e^{-(s-j\omega_0)T}}{s - j\omega_0}\right|_{s = j\omega_0} = T$$

Thus an impulse sampler attenuates a continuous harmonic input signal by a factor of "T". Such attenuation should be reflected in the frequency response curves; a fact witnessed by the example curves (solid lines) of figures (8a) and (8b) where T = 0.5. Since attenuation factor "T" is independent of forcing frequency w_o, dividing the input-output ratio by "T" would bring the zero frequency response to unity and the sampled data frequency response in general would closely correspond to that for continuous systems. These desirable ends supply the motive for the bandwidth definition of equation (55). Observe that the frequency response curves as defined for z-transfer functions

$$A = \left|\frac{C(z)}{R(z)}\right|_{z = e^{j\omega_0 T}} \equiv \left|\frac{C^*(s)}{R^*(s)}\right|_{s = j\omega_0}$$

[shown dashed in figures (8a) and (8b)] have this attenuation factor "T" present in both numerator and denominator. Hence these curves are conceptually satisfying without recourse to scaling. The curves of figures (9) are like those of figures (8) except that s-plane frequency response definitions are scaled to have unity value for inputs of $w_o = 0$ and Y-axis coordinates are in decibels. The parameter settings for two of the three curves shown

a) $k = 0.44$, $k_t = 0.177$ figure (9a)

b) $k = 0.39$, $k_t = 1.643$ figure (9b)

lie on the same parameter plane bandwidth curve in figure (10); consequently each frequency curve should exhibit a magnitude of 0.3 (-10.5 db) at frequency $w = \frac{\pi}{2}$. Although the frequency response curves of figures (9) apply to but three specific examples the reader should recognize the extreme low pass properties of the hold element and the close correspondence between

$$\frac{1}{T}\left|\frac{C(s)}{R^*(s)}\right| \quad \text{and} \quad \left|\frac{C^*(s)}{R^*(s)}\right|$$

for $w < \frac{\omega_s}{2}$ which these curves illustrate. These observations are generally valid although the frequency at which the generalizations cease to apply is dependent on the particular system and the sampling rate.

These digressions are relevant to parameter plane study to the extent that they allow meaningful specification and interpretation of bandwidth loci thereon. Several such loci are shown (solid curves) in figure (10). For the specified frequency each curve divides the parameter plane into regions having attenuation greater than or less than the attenuation evidenced by the parameter settings on the curve. If the curve itself represents points just meeting a bandwidth requirement then parameter settings corresponding to points in one of these regions will all exceed the bandwidth specification. Moreover no intersection of bandwidth loci occurs for various

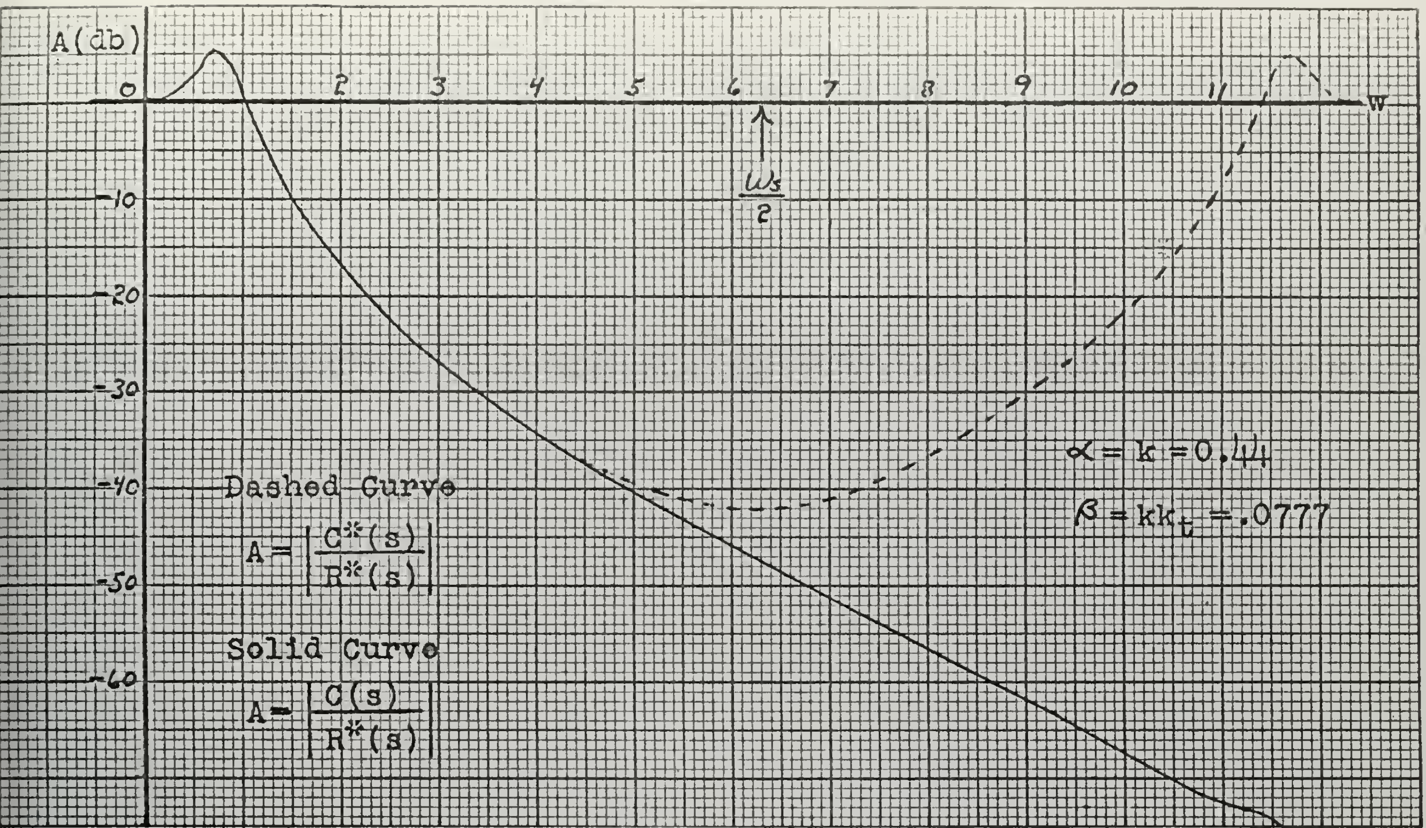


FIGURE 9a

FREQUENCY RESPONSE CURVES
FOR A SAMPLED DATA SYSTEM

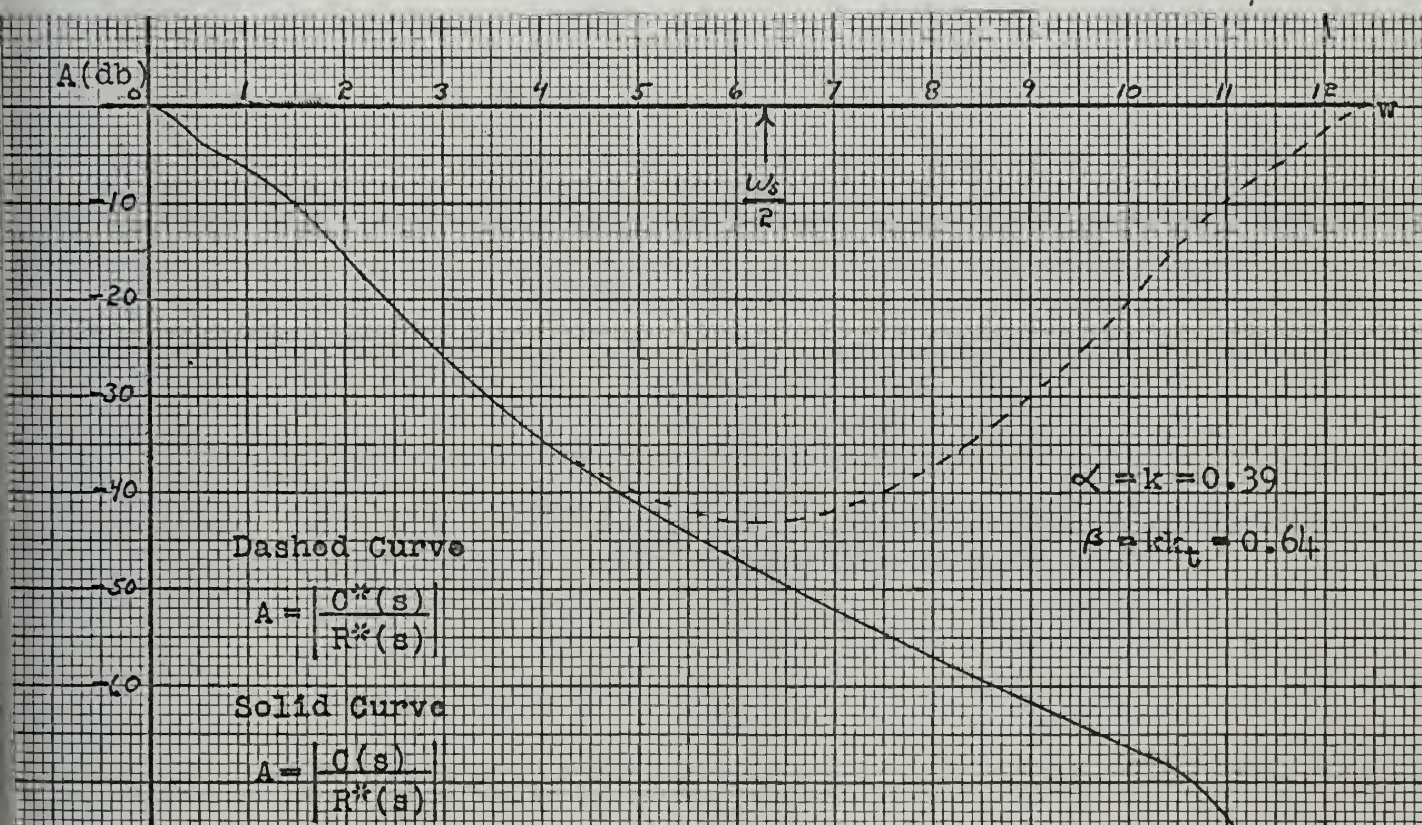


FIGURE 9b

FREQUENCY RESPONSE CURVES
FOR A SAMPLED DATA SYSTEM

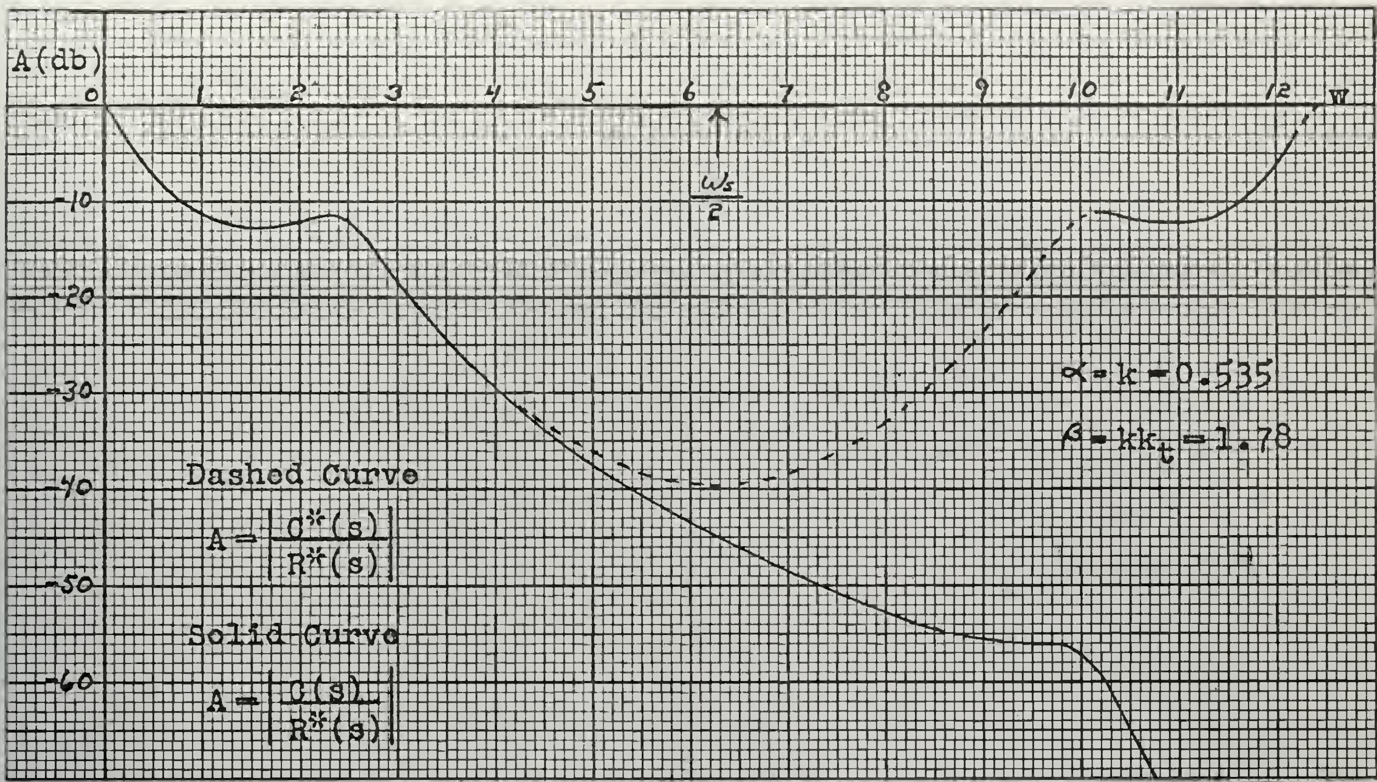


FIGURE 9c

FREQUENCY RESPONSE CURVES
FOR A SAMPLED DATA SYSTEM

The bandwidth and steady state error loci displayed in figure 10 are for the third order sampled system shown in figure 2. Equation 55 governing these bandwidth loci is:

$$A = \frac{1}{T} \left| \frac{\dfrac{\alpha(1-e^{-sT})}{s^2(s+1)(s+2)}}{1 + \dfrac{(.058\alpha + .31\beta)e^{2sT} + (.16\alpha - .12\beta)e^{sT} + (.026\alpha - .19\beta)}{e^{3sT} - 1.97e^{2sT} + 1.20e^{sT} - 0.22}} \right|$$

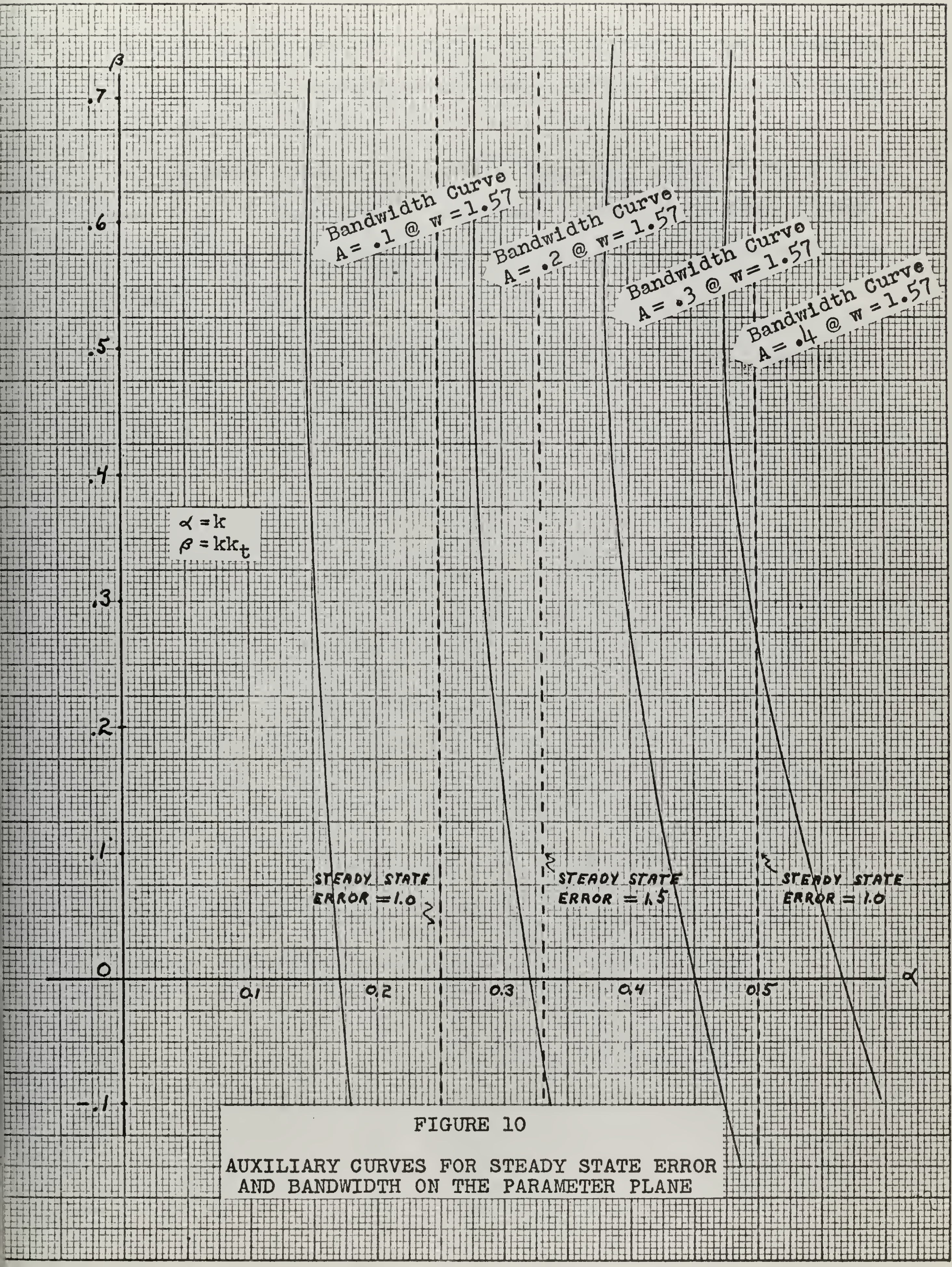


FIGURE 10

AUXILIARY CURVES FOR STEADY STATE ERROR AND BANDWIDTH ON THE PARAMETER PLANE

magnitude specifications at a given frequency although curves representing the same magnitude attenuation specified at various frequencies can conceivably have common points.

The dashed straight line curves in figure (10) represent steady state error loci. Perhaps their most significant feature is the vertical slope from which the reader infers that steady state error is independent of the parameter $\beta = kk_t$.

CHAPTER V

Concepts of Dominant Mode Design

The preponderance of system criteria in the s-plane are based on dominant mode analysis and design. Such concepts as damping factor, resonant frequency, bandwidth and maximum overshoot are meaningful only within the context of a dominant mode system. (i.e. What meaning has bandwidth if an attenuation of -3 db occurs at three points on the frequency responce curve?) Rudiments of dominant mode design reveal the direction any extensions of this method must take.

1. An ideal model must be defined. (For s-plane use, the no-zero, two complex pole model is universally accepted as standard).
2. The actual system must be compensated to behave like the ideal model. (In the s-plane the actual system approximates the ideal second order model if a pair of complex poles dominate its behavior).

Dominant mode design in the parameter plane follows the indicated path. As an initial step a z-plane "ideal model" is defined and system behavior (rise time, maximum overshoot) is related to model parameters. The following section discusses conditions requisite on other systems to insure dominant mode (or model-like) behavior. Finally the parameter plane is introduced as a vehicle for designing or compensating the actual system to meet the requirements for guaranteed dominant mode performance.

The Dominant Mode Sampled Data System Model and Its Specification

Choice of a model is critical because dominant mode design can never intentionally produce systems whose properties exceed those of the model. (To illustrate consider an s-plane design in which a fourth order system is

compensated to have two dominant complex poles. It may be that the fourth order system when compensated outperforms the second order model but such performance is coincidental rather than designed). An elaborate model might be justified were it not for two additional requirements.

1. Common specifications must be readily translatable as model parameters (root locations). By way of example, a maximum overshoot specification can be translated into a zeta restriction on the roots of the second order s-plane model.
2. The model's pole-zero configuration must have simplicity sufficient to enable the designer to recognize when model like behavior (dominance) is achieved. Suppose, for an example by contrast, it is possible to determine a fourth order ideal model system. A designer now compensates an existing system to approach this model; but approach it where? Should he design to have two poles in close proximity and two further removed or should all poles be moderately removed from those of the model or is some other approximation to be preferred?

These considerations indicate an acceptable model to be one whose closed loop z-transfer function is

$$\frac{C(z)}{R(z)} = \frac{k(z - z_1)}{(z - p_1)(z - \bar{p}_1)} \tag{58}$$

where z_1 is a real zero

p_1 and $\bar{p}_1$ are conjugate complex poles

$k = \frac{(1-p_1)(1-\bar{p}_1)}{(1-z_1)}$ This choice of "k" guarantees a type I system with zero steady state error to unit step inputs.

The utility of this model is contingent upon the facility with which system specifications are translatable into acceptable pole-zero configurations. To amplify, unless acceptable model pole-zero configurations can be simply determined design might profitably proceed in some direction other than in forcing the system to resemble a model. The following paragraphs develop the required relations between maximum overshoot, rise time and model pole-zero locations. This development largely duplicates the work on pages 170 through 179 in Kuo and is included to accentuate approximations inherent in the model itself [2]. These approximations are over and above those made when compensating higher order systems to a dominant mode and are not required for continuous systems.

The rise time and maximum overshoot of the system described by equation (58) to a unit step input are desired. This requires expression of c*(t).

$$C(z) = \frac{k(z - z_1)}{(z - p_1)(z - \bar{p}_1)} \cdot \frac{z}{z-1}$$

By the inversion integral:

$$c(nT) = \frac{1}{2\pi j} \int_{\Gamma} \frac{k(z - z_1)\, z}{(z-1)(z - p_1)(z - \bar{p}_1)}\, z^{N-1}\, dz$$

where Γ is a circle of radius e^{rT} centered at the origin and enclosing all poles of C(z). Applying the residue theorem to this integral one obtains

$$C(nT) = \frac{k(1 - z_1)}{(1 - p_1)(1 - \bar{p}_1)} + \frac{k(p_1 - z_1)\, p_1^N}{(p_1 - 1)(p_1 - \bar{p}_1)} + \frac{k(\bar{p}_1 - z_1)\, \bar{p}_1^N}{(\bar{p}_1 - 1)(\bar{p}_1 - p_1)}$$

which under definition of

$$p_1 = |P|\, e^{j\Phi}$$

becomes

$$c(nT) = 1 - 2k \frac{|(p - z_1)\, p^N|}{|(p - 1)(p_1 - \bar{p}_1)|} \sin(\gamma + N\Phi)$$

or

$$c(nT) = 1 + 2k\,\frac{|(p-z_1)\,p^N|}{|(p-1)(p_1-\bar{p}_1)|}\;\cos\,\theta_1 + n\Phi$$

where $\gamma = \arg(p_1-z_1) - \arg(p_1-1)$

$$\theta_1 = \arg(p_1-z_1) - \arg(p_1-1) - \frac{\pi}{2}$$

With specification of the closed loop poles and zero the above expression is sufficient to determine the output c*(t) at the sampling instants. However this expression has two serious defects.

1. The output is specified only at the sampling instants whereas peak overshoot in the continuous output is likely to occur during the intersampling period.

2. The output c*(t) is specified as a function of three parameters, a) real part of complex poles, b) imaginary part of complex poles and c) location of real zero; consequently its universal solution is difficult to display. That is to say - system criteria which are functions of three variables are not conveniently represented in two dimensions or less.

The next step is circumvention of these difficulties. The first problem can be attacked in several ways among which are employment of modified z-transforms and approximation of the continuous output c(t) by a continuous function $c_a(t)$ which agrees with c(t) at the sampling instants. The complexity of the modified z-transform and the fact that zeros of C(z,m) are functions of m preclude the use of this avenue. Approximation of the continuous output with a continuous function which agrees at the sampling instants offers many advantages. Among them, the peak overshoot of this approximating function can be determined and if, as is hoped, the approximation is good then this peak should be close to the true peak overshoot. The second problem is somewhat more difficult. It involves specifying the poles and zeros

in such a manner that overshoot can be expressed in terms of two parameters instead of three. Offhand there is no direct procedure but rather it is one of trial and error. The following examples should clarify what is being attempted.

1. For continuous second order systems the complex poles can be expressed as $\sigma + jw$ or $w_n(\zeta + j\sqrt{1-\zeta^2})$. When the poles are expressed by ζ and w_n then overshoot becomes a function of ζ alone.

2. Suppose $f(\alpha, \beta, \gamma) = \alpha^2\beta^2 - \alpha\beta + \gamma$. It is obvious that f can be expressed as a function of two variables by defining $\delta = \alpha\beta$. Then

$$f = (\alpha, \beta, \gamma) = \delta^2 - \delta + \gamma$$

For the second order sampled data system of interest a successful choice of variables has been found. Define:

$$P_1 = e^{-\zeta \omega_n T} e^{j\sqrt{1-\zeta^2}\,\omega_n T}$$

$$\delta = \mathrm{ARG}(P_1 - Z_1) - \mathrm{ARG}(P_1 - 1) + \frac{\pi}{2}$$

Expression of c(nT) in terms of these parameters is most easily accomplished by recourse to geometry.

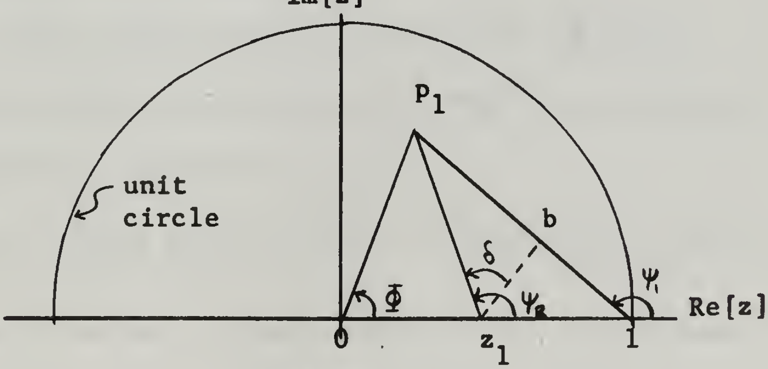


Figure 11

Consider figure 11. By inspection the following definitions can be made.

Ψ_1 = argument $(p_1 - 1)$

Ψ_2 = argument $(p_1 - z_1)$

Φ = argument (p_1)

$\delta = \arg(p_1 - z_1) - \arg(p_1 - 1) + \frac{\pi}{2} \equiv \Psi_2 - \Psi_1 + \frac{\pi}{2}$

These definitions of δ and Φ are consistent with those made earlier. An interesting representation for the secant of δ results from these definitions. Define the secant in terms of the legs of right triangle (p_1, z_1, b).

$$\left|\sec \delta\right| = \left|\frac{P_1 - Z_1}{Z_1 - b}\right|$$

Now observe that the area of triangle (p_1, z_1, 1) can be expressed in two ways

$$\text{area } \Delta(p_1, z_1, 1) = \frac{1}{2}\left(\frac{1}{2}|P_1 - \bar{P}_1||1 - Z_1|\right) = \frac{1}{2}\left(|Z_1 - b||1 - P_1|\right)$$

$$|z_1 - b| = \frac{1}{2}\frac{|P_1 - \bar{P}_1||1 - Z_1|}{|1 - P_1|}$$

Therefore

$$\left|\sec \delta\right| = 2\frac{|1 - P_1||P_1 - Z_1|}{|P_1 - \bar{P}_1||1 - Z_1|} = 2\frac{|1 - P_1||1 - \bar{P}_1|}{|1 - Z_1|}\frac{|P_1 - Z_1|}{|P_1 - \bar{P}_1||1 - \bar{P}_1|}$$

$$\left|\sec \delta\right| = 2k\frac{|P_1 - Z_1|}{|P_1 - \bar{P}_1||1 - \bar{P}_1|} = 2k\left|\frac{(P_1 - Z_1)}{(P_1 - 1)(P_1 - \bar{P}_1)}\right|$$

But this expression is identical to that which appears in the equation for c(nT). Hence c(nT) can be written as:

$$c(nT) = \left|\sec \delta\right| |P_1|^N \cos\left(N\Phi + \delta - \pi\right)$$

The previously defined representation of "p_1" as a function of ζ and w_n allows expression of c(nT) as:

$$c(nT) = \left|\sec \delta\right| e^{-\zeta \omega_n nT} \cos\left(\sqrt{1 - \zeta^2}\, \omega_n nT + \delta - \pi\right)$$

At this point it becomes a simple matter to fit a continuous curve through the discrete output c(nT). To do this simply substitute the continuous

variable t for the discrete variable nT.

$$c_a(t) = |\sec \delta|\, e^{-\zeta \omega_N t} \cos\left(\sqrt{1-\zeta^2}\,\omega_N t + \delta - \pi\right) \tag{59}$$

$c_a(t)$ is the continuous function which will be used in place of the true output c(t). While $c_a(t)$ coincides with c(t) at sampling instants there is no requirement that it coincide during the intersampling periods as well. Nevertheless it is assumed that $c_a(t)$ is "close to" c(t) at all times and that the system output is adequately represented by $c_a(t)$. Under this assumption $c_a(t)$ can be differentiated to obtain peak time, maximum overshoot and other characteristics inherent in c(t).

The peak time is found by differentiating $c_a(t)$ with respect to t and equating the resulting expression to zero.

$$T_{max} = \frac{1}{\omega_N \sqrt{1-\zeta^2}} \left[\tan^{-1}\left(\frac{-\zeta}{\sqrt{1-\zeta^2}}\right) - \delta + \pi \right] \tag{60}$$

The maximum overshoot is found by substituting the expression for T_{max} into $c_a(t)$.

$$c_a(\max) = \underline{\sqrt{1-\zeta^2}\, |\sec \delta|\, e^{-\frac{\zeta}{\sqrt{1-\zeta^2}}\left[\tan^{-1}\left(\frac{-\zeta}{\sqrt{1-\zeta^2}}\right) - \delta + \pi\right]}} + 1 \tag{61}$$

The portion of this expression underlined is the maximum overshoot to a unit step input. Charts of overshoot and $\frac{T_{MAX}\,\Phi}{T}$ as functions of δ and ζ can be found on pages 175 and 176 of Kuo [2].

The equations for T_{max} (60) and $c_a(\max)$, (61) characterize a second order sampled data system in terms of two parameters. However they are not exact representations and the problem is not really complete until regions of validity for this characterization are determined. Lindorff (by what seems to be intuitive argument) has determined that the approximation is good if the complex poles are in the first and fourth quadrants and are not too close to the origin of the z-plane [9]. [Poles in the first and fourth

quadrants insure at least four samples per cycle of the dominant oscillatory mode and thus give a reasonably good approximation for this mode. However, if the poles are too close to the origin the rapid decay offered by the $|p|^n$ factor will predominate to the point where the z-transform [and thus the approximating function $c_a(t)$] no longer adequately characterize the system. That is, the sampling rate is too slow to follow the decay].

Analysis of Dominance Requirements

For the z-plane dominant mode model developed in Kuo and reviewed in the preceding section to have value the conditions under which model-like behavior prevails must be determined. These requirements can be established by considering the inverse z-transform integral.

$$c(nT) = \frac{1}{2\pi j}\int_{\Gamma} C(z)\, z^{N-1}\, dz$$

where Γ is a path which encloses all singularities of C(z). But

$$C(z) = G(z)R(z)$$

$$C(z) = \frac{\prod^{M}(z-z_k)}{\prod^{N}_{\ell=1}(z-p_\ell)}\ \frac{z}{z-1} \qquad \text{for a unit step input.}$$

Substituting this expression into the integral equation yields

$$c(nT) = \frac{1}{2\pi j}\int \frac{\prod_{k=1}^{M}(z-z_k)}{\prod_{\ell=1}^{N}(z-p_\ell)}\ \frac{z}{z-1}\ z^N\, dz$$

This integral can now be evaluated by the theory of residues. Expressed in this manner it becomes:

$$c(nT) = \sum_{\lambda=1}^{N}\frac{p_\lambda^N}{(p_\lambda-1)}\ \frac{\prod_{k=1}^{M}(p_\lambda-z_k)}{\prod_{\ell=1,\ \ell\neq\lambda}^{N}(p_\lambda-p_\ell)} + \frac{\prod_{k=1}^{M}(1-z_k)}{\prod_{k=1}^{N}(1-p_k)} \leftarrow \text{steady state output}$$

Analysis of this expression reveals the two conditions necessary for dominant mode behavior.

1. All but the two dominant poles must lie near the origin of the z-plane. After a small number of samples the residues

contributed by poles near the origin are negligible; the p^n term dominates these residues and as a factor it becomes almost zero.

2. Excess zeros over the real zero employed in the second order model must lie near the origin of the z-plane. When so located two effects are assumed to result.

   a. The residues of poles near the origin (which are neglected in the second order model) will be small.

   b. These zeros form dipoles (pole-zero pairs close together) which have almost negligible effect on the residue of the dominant poles. [i.e. If $(p_i-z_k) \simeq (p_i-p_k)$ then these common factors cancel in the expression for the dominant pole p_i]. (Where time delays are absent, dominant mode design requires as many poles as zeros to be ignored or lie near the origin. This proximity of ignored poles and ignored zeros essentially guarantees dipole formation when viewed from the location of the dominant poles).

Principles of Dominant Mode Compensation in the Parameter Plane

With model behavior and restrictions to insure model behavior in higher order system established attention turns to the problem of properly compensating existing systems to resemble the model. Parameter plane methods are well suited to this aspect of dominant mode design. This approach not only allows selection of the dominant (model) poles but also guarantees dominant mode (or model like) behavior by forcing compliance with the restrictions mentioned above. Tacit is the assumption that parameter values (α, β) exist which allow accomplishment of this ambitious objective. Justification of these

statements can be had by recourse to Chapter III. Recall that the image of a z-plane contour divides the parameter plane into regions whose distinguishing characteristic is that any parameter setting within a given region places an equal number of characteristic equation roots within the z-plane contour which was mapped. Suppose a circular contour centered at the origin of the z-plane and having a radius small enough so that all poles within it can be considered negligible, is mapped through the characteristic equation into the parameter plane. Of the parameter plane regions defined by this mapping the dominant mode designer considers only those for which all roots but two of the characteristic equation lie within the z-plane contour. If within such a region parameter settings exist which place two poles near the unit circle then these poles must be dominant (because it is known that no other poles except those near the origin exist). To find these poles auxiliary contours of constant w_n and constant ζ are mapped through the characteristic equation into the parameter plane. The intersection of an acceptable w_n contour and ζ contour in the region mentioned above constitutes a set of parameter values which simultaneously guarantees dominance and the desired pole locations.

To this point the closed loop zero specification required by the dominant mode model has been ignored. Such slight is intentional since zeros of the open and closed loop transfer function are identical and are not generally subject to modification by available system parameters. Therefore discounting cancellation, zeros should be considered as entities which allow or disallow dominant mode compensation rather than as entities to be modified.

Dominant mode design using parameter plane concepts is readily extended to continuous systems described in the s-plane. Instead of a small circle near the origin a contour of constant σ ($\sigma = \zeta w_n$) sufficiently removed from

the imaginary axis is mapped through the characteristic equation. The image of this s-plane contour divides the parameter plane in exactly the same manner as the image of the small circle in the z-plane; subsequent s and z-plane procedures are identical. In fact dominant mode design in the s-plane is generally simpler than in the z-plane because the s-plane dominant mode model has no zeros in the numerator of its transfer function.

Cascade Compensation (for Dominant Mode Design)

The principles discussed in the preceding chapter fairly well dictate the form that a cascade compensator must have if dominant mode design is to be achieved. Consider the sampled data system of figure (12).

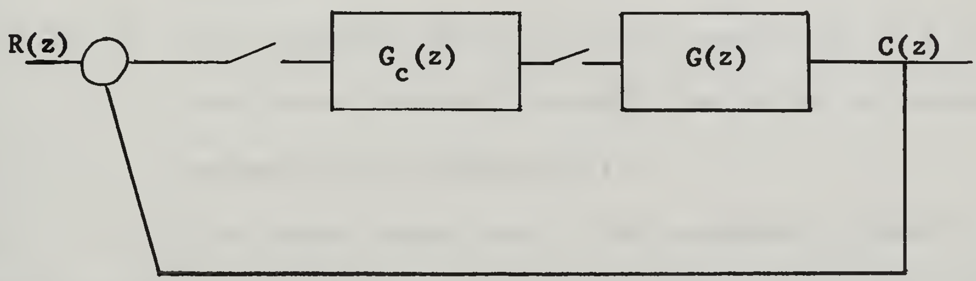


Figure 12

G(z) is the plant transfer function which is assumed to be of the form:

$$G(z) = \frac{k \prod_{\lambda=1}^{M} (z + z_\lambda)}{\prod_{\lambda=1}^{N} (z + P_\lambda)} \qquad (62)$$

The system open loop zeros provide the best clue as to the feasibility of dominant mode design. Since the zeros of the open and closed loop transfer functions are identical, all zeros except one must lie near the origin of the z-plane. When this condition is not met no cascade compensation is capable of driving the system to the second order dominant mode desired. It is also prudent to observe the system poles before undertaking dominant mode design. Cascade compensation can succeed, disallowing pole-zero cancellation, only if addition of poles and zeros offers the possibility of

reshaping the root locus so that all but two complex poles and one real zero lie near the origin. If the general uncompensated system pole-zero configuration is such that this is manifestly impossible then there is no point in employing parameter plane procedures to prove it.

These observations can be made more rigorous by translating them into restriction on the transfer function of the cascade compensator.

1. The zeros of the cascade compensator must lie near the origin of the z-plane (assuming that one system zero is not so located). In this way these zeros can be neglected without affecting the validity of the dominant second order approximation.
2. The numerator should be of the same order in z as the denominator (primarily because time delay is an undesirable property in a compensator).
3. The steady state gain of the compensator should be unity.

With these properties in mind the transfer function of the cascade compensator must have the form:

$$G_c(z) = \frac{(1 + a_{N-1} + \cdots + a_1 + a_0)\, Z^N}{Z^N + a_{N-1} Z^{N-1} + \cdots + a_1 Z + a_0} \tag{63}$$

With this cascade compensator installed the closed loop transfer function of the system in figure (12) is expressible as:

$$\frac{C(z)}{R(z)} = \frac{k\,(1 + a_{N-1} + \cdots + a_1 + a_0)\, Z^N \prod_{i=1}^{M} (Z - Z_i)}{(Z^N + a_{N-1} Z^{N-1} + \cdots + a_0) \prod_{i=1}^{E} (Z - P_i) + k\,(1 + a_{N-1} + \cdots + a_0)\, Z^N \prod_{i=1}^{M} (Z - Z_i)} \tag{64}$$

where k, z_i, and p_i are known fixed parameters of the plant. The a_j's are now chosen to yield a satisfactory dominant mode second order system. While equations (63) and (64) appear formidable, parameter plane methods will yield

a solution almost by inspection especially if but two of the a_j's are variable.

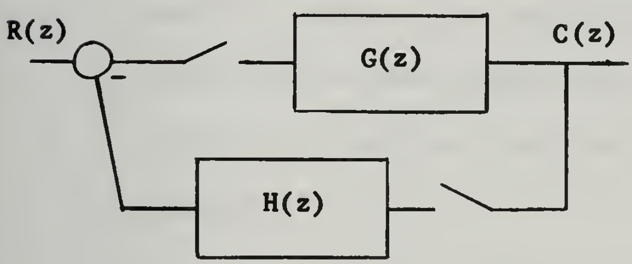
This method is also applicable when pole-zero cancellation is allowed. In this case the cascade compensator has a transfer function of the form:

$$G_c(z) = \frac{(1 + a_{N-1} + \cdots + a_1 + a_0) Z^N \prod_{i=1}^{m} (Z + r_i)}{(Z^N + a_{N-1} Z^{N-1} + \cdots + a_1 + a_0) \prod_{i=1}^{r} (Z + t_i)}$$

where $(z + r_i)$ and $(z + t_i)$ cancel poles and zeros of the plant transfer function respectively.

Feedback Compensation (for Dominant Mode Design)

Feedback compensation instead of cascade compensation can be used to effect dominant mode design. As with cascade compensation certain restrictions are placed on the feedback transfer function if the chosen second order dominant model is to be valid. Consider the sampled data system of figure (13).



where H^n = numerator of H(z)

H^d = denominator of H(z)

G^n = numerator of G(z)

G^d = denominator of G(z)

Figure (13)

Assume that:

a. The plant transfer function G(s) contains at least one integration. (i.e. The denominator of G(s) contains a factor of s or the denominator of G(z) contains a factor [z-1[).

b. At most all but one zero of the plant transfer function is located near the origin of the z-plane.

Requirements:

1. Design a feedback compensator H(z) which yields a second order dominant mode system with acceptable time response characteristics.

2. The system must follow a step input with zero steady state error.

The first step toward solution will be to apply the assumptions with the objective of obtaining an acceptable class of transfer functions H(z). Parameter plane analysis applied to this class will then yield acceptable H(z)'s if they exist. The closed loop transfer function for the system of figure (13) is

$$\frac{C(z)}{R(z)} = \frac{G(z)}{1 + G(z)H(z)}$$

$$\frac{C(z)}{R(z)} = \frac{G^N H^d}{G^d H^d + G^N H^N}$$

The first restriction on H(z) can be obtained by observing that poles of H(z) are zeros of the closed loop transfer function. To effect dominant mode design these closed loop zeros [or equivalently poles of H(z)] must lie near the origin. The second restriction on H(z) results from the requirement that there be zero steady state error to a step input. This restriction is obtained by applying the final value theorem to the closed loop transfer function.

$$c(\infty) = \lim_{z\to 1} \left[\frac{(z-1)}{z} R(z) \frac{G^N H^d}{G^d H^d + G^N H^N}\right] = 1 \qquad \text{for unit step input}$$

$$c(\infty) = \lim_{z\to 1} \left[\frac{(z-1)}{z} \frac{z}{(z-1)} \frac{G^N H^d}{G^d H^d + G^N H^N}\right] = \left[\frac{G^N H^d}{G^d H^d + G^N H^N}\right]$$

But $\lim_{z\to 1} [G^d] = 0$ since the plant transfer function was assumed to contain an integration. Hence:

$$c(\infty) = \lim_{z\to 1} \left[\frac{H^d G^N}{H^N G^N}\right] = \lim_{z\to 1} \left[\frac{H^d}{H^N}\right] = 1$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

or

$$\lim_{z \to 1} H(z) = 1$$

To meet these requirements the feedback path must have a transfer function of the form

$$H(z) = 1 + \frac{(z-1)\sum_{k=0}^{N} a_k z^k}{z^{N+1}} \tag{65}$$

This transfer function is slightly more restrictive than necessary since all poles of H(z) were placed at the origin (for computational ease) rather than near the origin which is the specific requirement. Also notice that H(z) is written as containing no time delays although this is not a specific requirement. Time delays can easily be added by increasing the powers of z in the denominator of H(z).

Example of Dominant Mode Design in the Parameter Plane

A numerical example is appropriate to illustrate theory and to reinforce stated results. Suppose the plant shown in figure 14 is to be compensated to a dominant mode which satisfies the specifications

$1.2 < M_{pt} < 1.4$

rise time - less than 5 sampling periods

As an initial step in the compensation procedures the zeros of this plant's closed loop transfer function (equation 66) are checked for compliance with necessary conditions for dominant mode design. Namely, validity of dominant mode design presupposes all zeros but one are sufficiently near the origin to be negligible. The real zero at z = -.295 stretches this requirement but when compared to the model zero at z = -3.72 it will probably not be sufficient to destroy dominant mode behavior. Tables describing model behavior such as those referenced in the first section of this chapter are now consulted for δ, ζ values which satisfy the specifications. For this

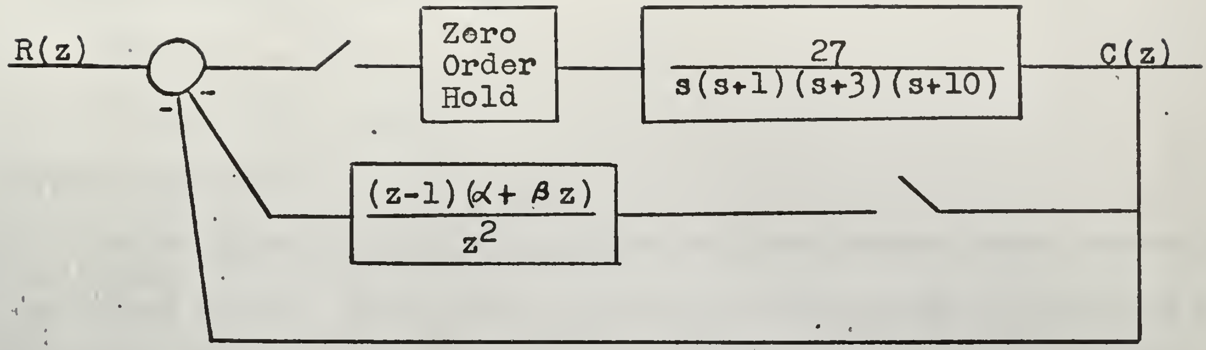


FIGURE 14

BLOCK DIAGRAM OF A FOURTH ORDER PLANT AND PROSPECTIVE COMPENSATION EMPLOYED TO EFFECT DESIRED DOMINANT MODE SYSTEM BEHAVIOR

The closed loop transfer function of the system shown above is:

$$\frac{C(z)}{R(z)} = \frac{N(z)}{D(z)}$$

where

$$N(z) = 0.022(z^2)(z+3.72)(z+0.295)(z+0.015)$$

$$\begin{aligned} D(z) = z^6 &+ (0.0220\beta - 1.814)z^5 \\ &+ (0.0220\alpha + 0.0667\beta + 1.066)z^4 \\ &+ (0.0667\alpha - 0.0632\beta - 0.116)z^3 \\ &+ (-0.0632\alpha - 0.0252\beta + 0.00128)z^2 \\ &+ (-0.0252\alpha - 0.00037\beta)z \\ &+ (-0.00037\alpha) \end{aligned} \tag{66}$$

example approximate value ranges

$$-30^\circ < \delta < 10^\circ$$

$$0.3 \leq \zeta \leq 0.45$$

appear satisfactory.

The parameter plane is now employed to effect dominant mode behavior in the actual system. Those system poles whose settling time is five times or more faster than that of the dominant mode will be considered negligible. Since w_{n_z} (the settling time indicator) of the dominant mode is about unity all other poles must lie within a circle of radius $w_{n_z} = 0.2$ centered at the origin. The image of this settling time contour divides the parameter plane into regions containing an equal number of roots to the characteristic equation (figure 15); the region for which all but two characteristic roots lie within the designated settling time contour is shaded. The presence of constant ζ curves in this region within the range needed to meet specifications is ascertained and on each curve the range of satisfactory δ values encountered in traversing the region is marked. These actions serve to delineate a parameter plane region (shaded in figure 16) within which all parameter settings provide a second order dominant system meeting specifications. Figure 17 compares the transient response of a true second order system with the response of the system in figure 14 compensated for dominant mode behavior. The dominant poles and zeros in the latter case identify with the poles and zero of the true second order system whose response is displayed. Numerical values for the compensated system are as follows:

parameter settings

$\alpha = -0.13$ $\qquad\qquad$ $\beta = 0.73$

dominant closed loop poles and zero

poles: $z = 0.776 \pm j0.362$ $\qquad\qquad$ zero: $= z = -3.72$

ignored closed loop poles and zeros

poles:	$z = 0.182 \pm j0.078$	zeros:	$z = -0.295$
	$z = -0.016$		$z = -0.015$
	$z = -0.106$		$z^2 = 0.0$

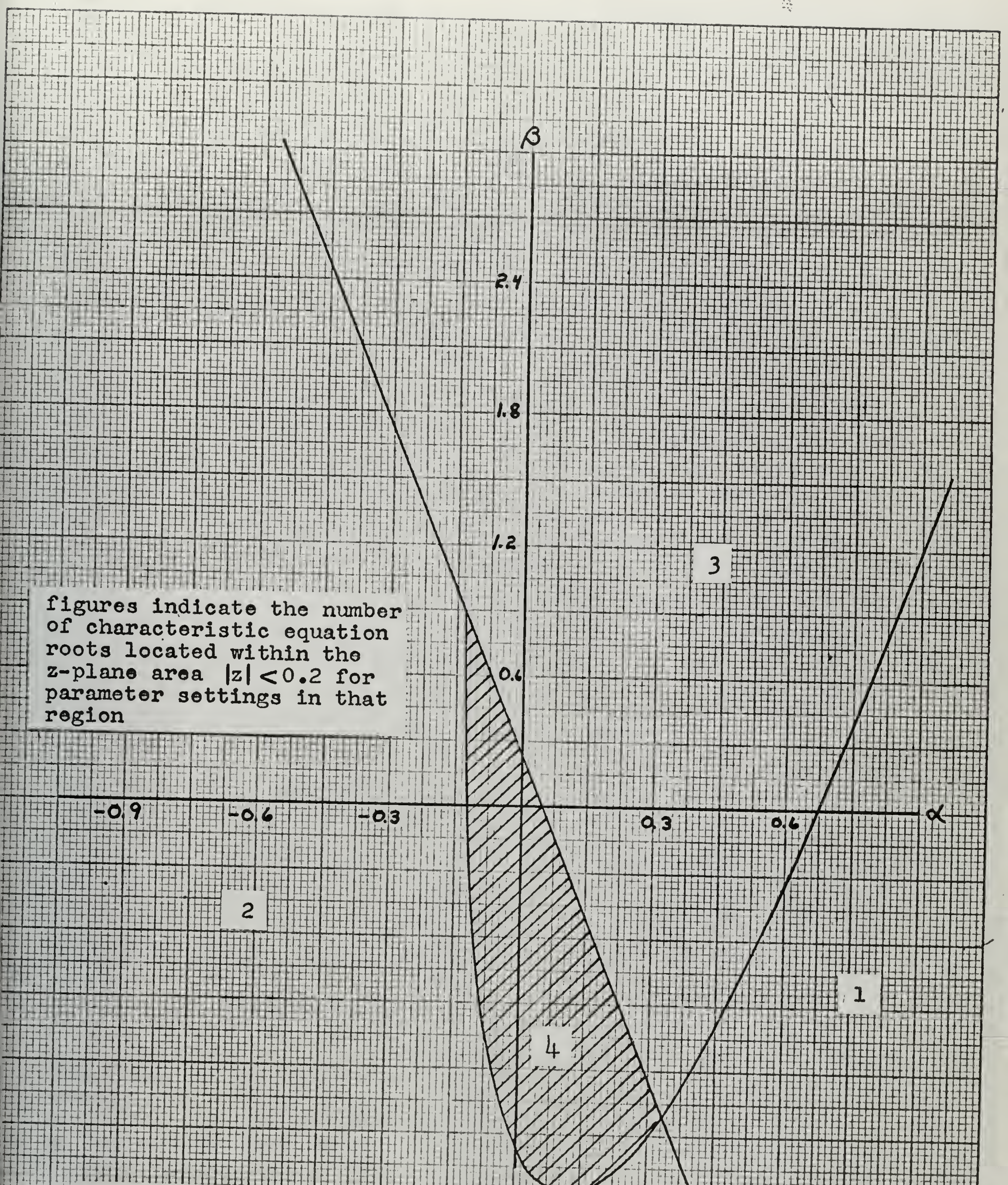


FIGURE 15

DIVISION OF THE PARAMETER PLANE INTO REGIONS BY THE IMAGE OF THE Z-PLANE CONTOUR $|Z| = 0.2$ PARAMETER SETTINGS WITHIN A GIVEN REGION LOCATE AN IDENTICAL NUMBER OF CHARACTERISTIC EQUATION ROOTS WITHIN THE Z-PLANE AREA $|Z| < .2$

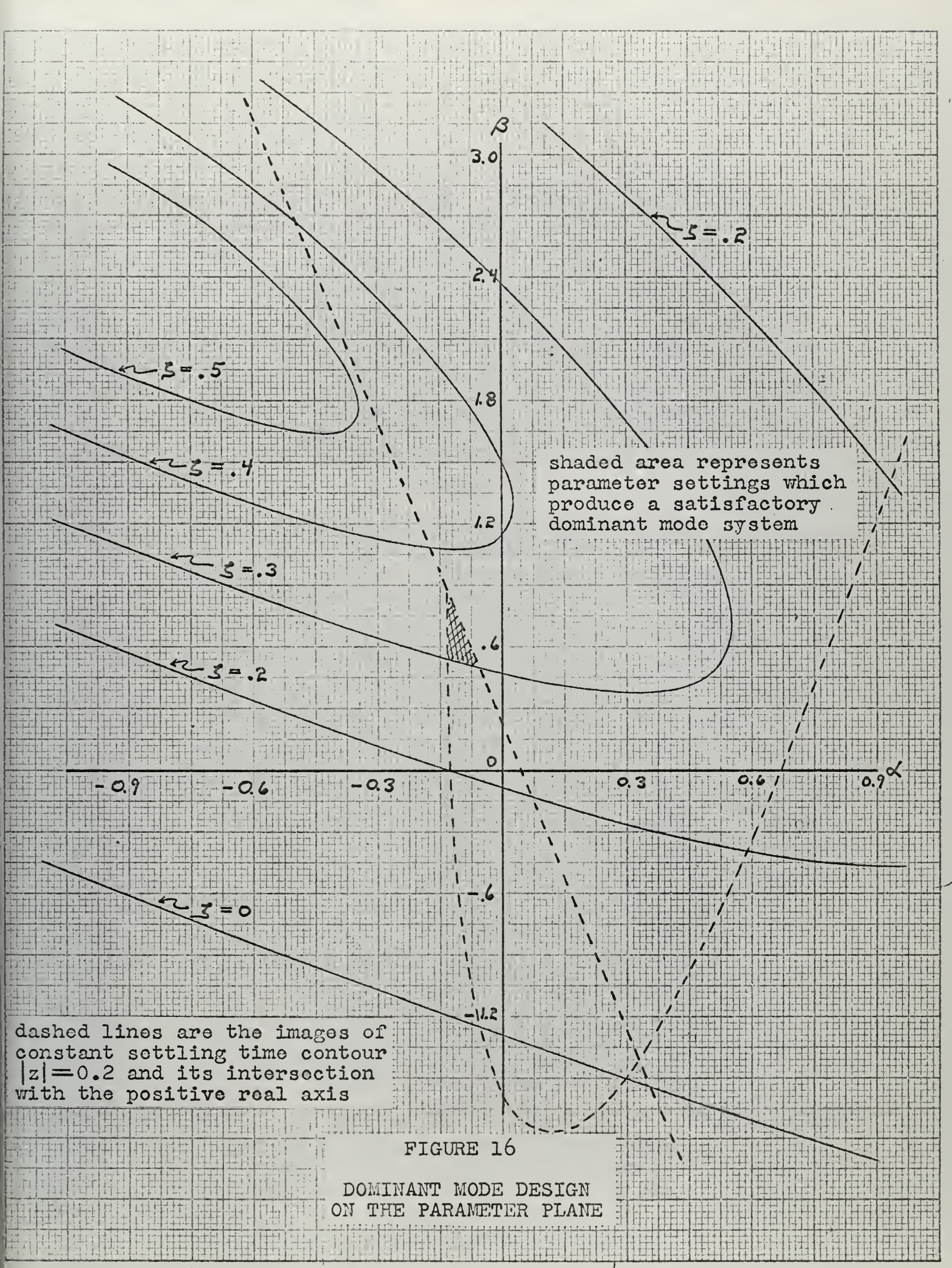


FIGURE 16

DOMINANT MODE DESIGN ON THE PARAMETER PLANE

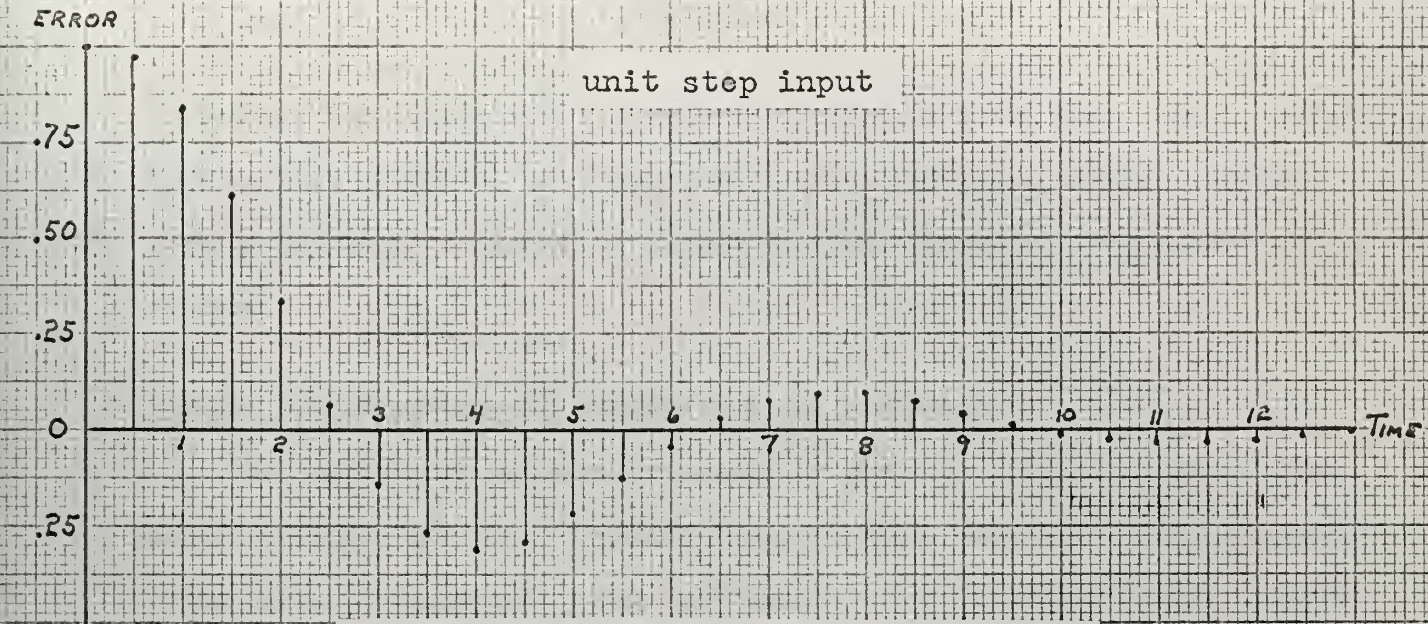


FIGURE 17a

TIME RESPONSE OF SIXTH ORDER SYSTEM
COMPENSATED FOR DOMINANT MODE BEHAVIOR

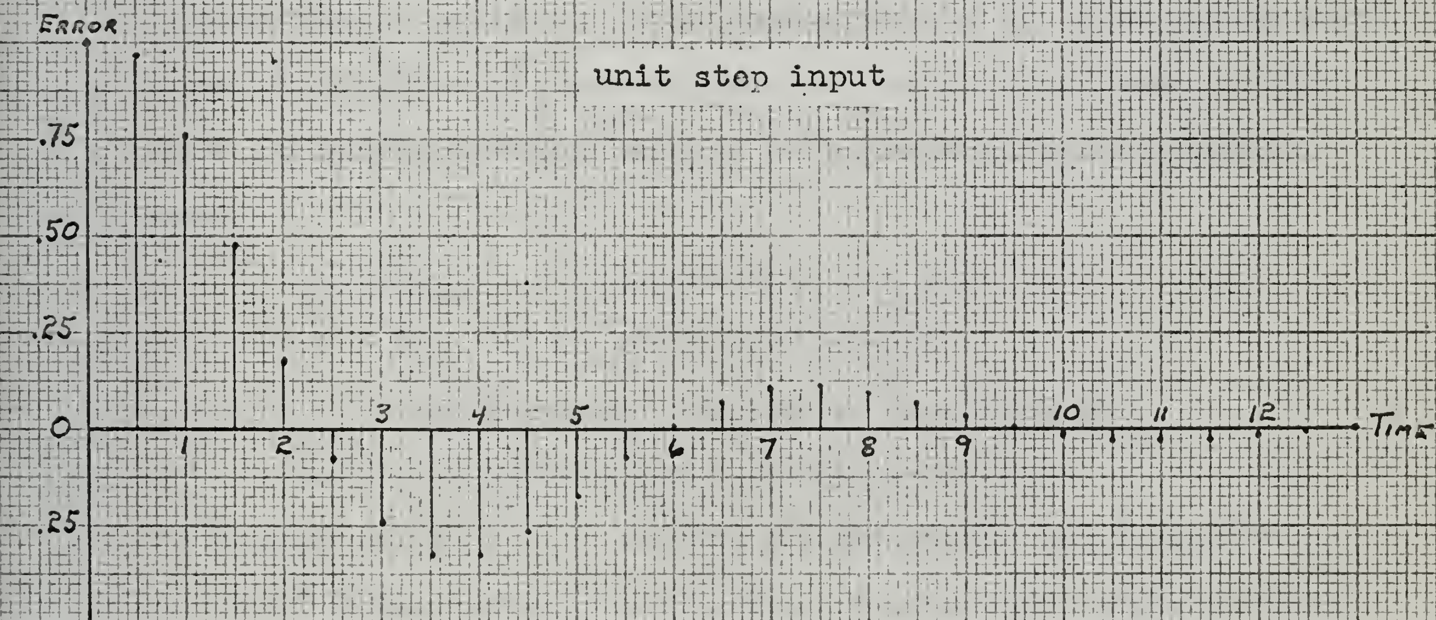


FIGURE 17b

TIME RESPONSE OF SECOND ORDER MODEL
WHOSE POLES AND ZERO CORRESPOND TO
THE DOMINANT POLES AND ZERO OF THE
SIXTH ORDER SYSTEM

CHAPTER VI

Frequently the niceties of a compensation procedure are more clearly revealed by numerical examples. In deference to this hypothesis several examples emphasizing various aspects of parameter plane theory are presented in this chapter.

Example of Compensation by Parameter Plane Technique

The fragmentary examples in preceding chapters relating to the system shown in figure (2) can be gathered into a fairly detailed and representative illustration of parameter plane technique. Typically the engineer is given a basic plant such as that shown in figure (2) and he is told to modify its performance to meet specification. Suppose these specifications require maximum overshoot to step inputs to be under twenty per cent, limit the sampling rate to two samples per second, and require steady state errors to unit ramp inputs of less than one. In addition short rise time and short settling time is desirable. Direct consideration of all these specifications is not practical in the parameter plane but tractable indicators for each exist. These specifications modified for parameter plane utilization become:

1. $\zeta_{S\text{-}PLANE} > 0.5$ in lieu of $M_{pt} < 1.2$
2. $E_{s.s.} < 1$ direct consideration
3. large bandwidth in lieu of short rise time
4. short settlint time direct consideration

The engineer decides that if feasible a combination of gain adjustment and derivative feedback offers the most attractive way to meet these specifications. The closed loop z-transfer function with the perspective compensation incorporated (equation 32) is amenable to parameter plane analysis.

[illegible]

Pertinent curves sufficient to determine the parameter settings required to effect a satisfactory system are shown in figure 18. One observes that to meet overshoot specifications parameter settings (α, β) must lie to the left of the $\zeta = 0.5$ contour while steady state error requires these settings to be to the right of the error locus shown. Unfortunately the intersection of these two areas is rather small leaving the designer little room to maneuver for the desirable properties of short settling time and short rise time. Nevertheless a constant settling time curve is shown; all parameter settings within the contour yield a system with a shorter settling time than is obtained for parameter settings outside the contour. Bandwidth curves which are indicative of rise time are also drawn; regions to the right of each curve have shorter rise times than those to the left. From these parameter plane considerations one concludes that settings of

$$\alpha = k \quad = 0.5$$

$$\beta = kk_t \quad = 0.5$$

should yield an acceptable system. The time response (at the sampling instants) for this choice of parameters is displayed for step and ramp inputs in figure 19a and 19b respectively.

Example of Analysis by Parameter Plane Technique

In addition to its capabilities as a design tool the parameter plane also provides an excellent vehicle for analysis. Consider the second and third order feedback systems shown in figure 20a and 20 b respectively and suppose them to be components of a still larger system. If pole b is far removed from pole a along the negative real axis it is well known that the third order system will behave essentially like its second order counterpart. However approximate second order behavior of this component depends not only

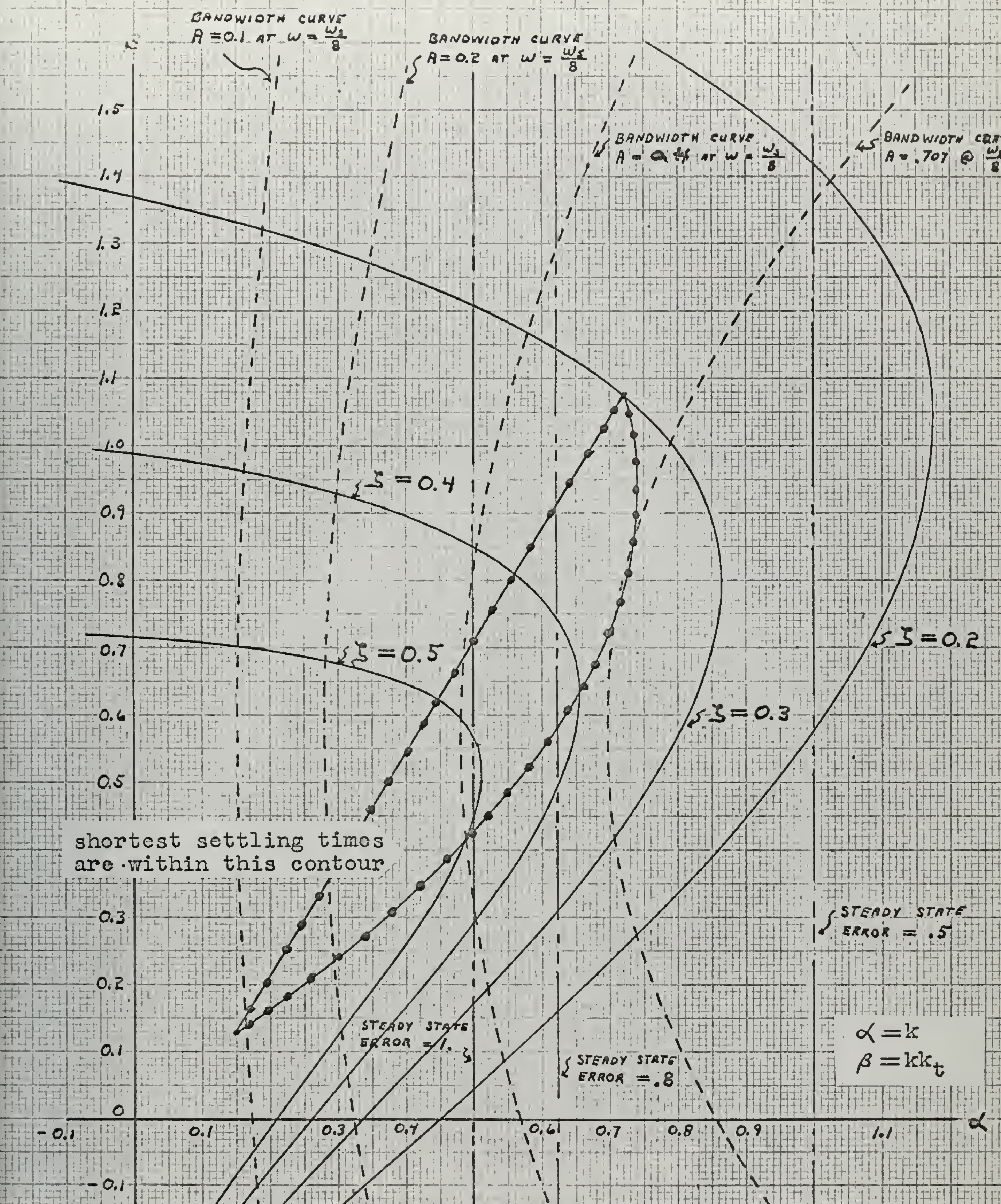
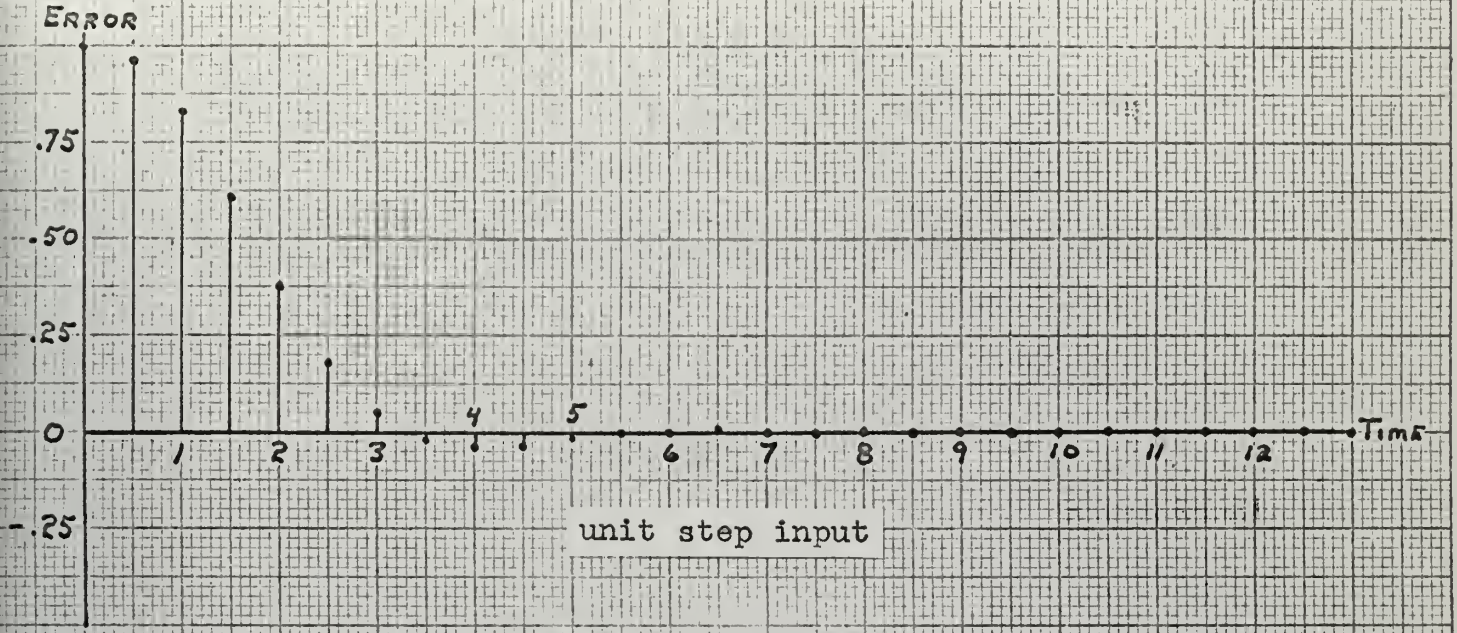


FIGURE 18

DESIGN OF COMPENSATION BY PARAMETER PLANE METHODS



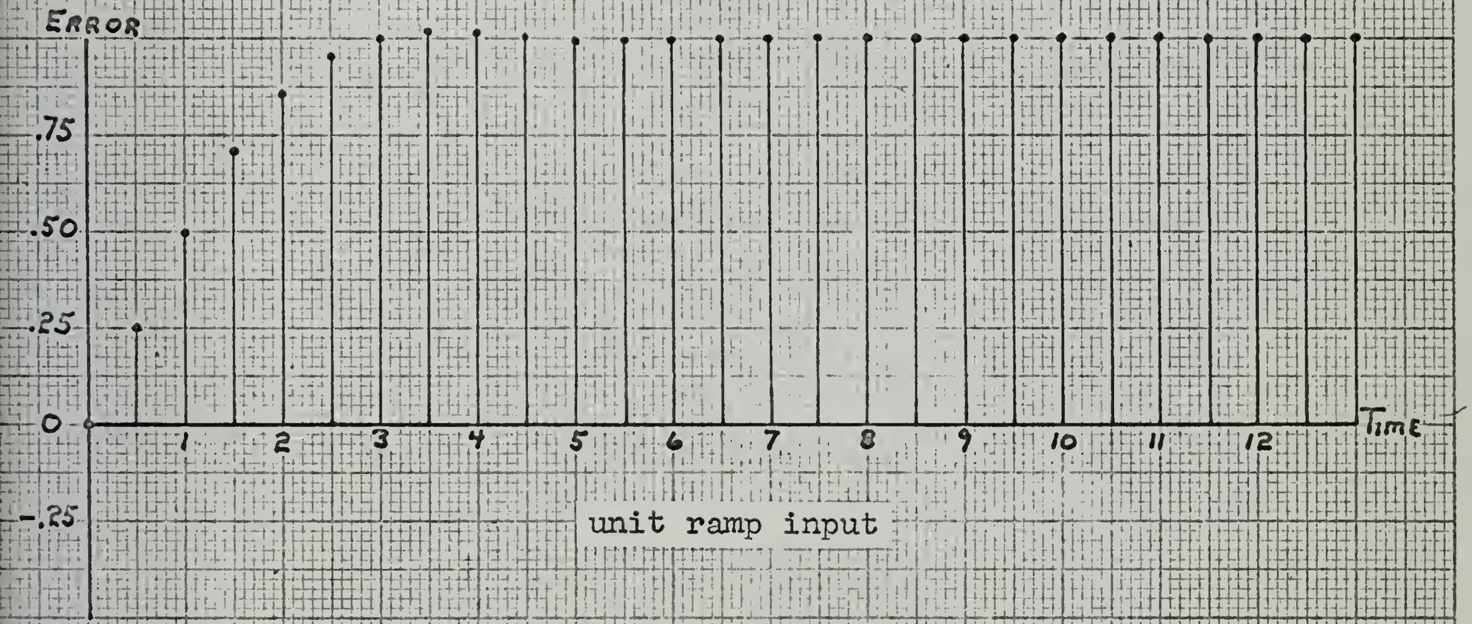


FIGURE 19

ERROR-TIME RESPONSE CURVES
FOR SYSTEM OF FIGURE 2
$\alpha = 0.5 \quad \beta = 0.5 \quad T = 0.5$

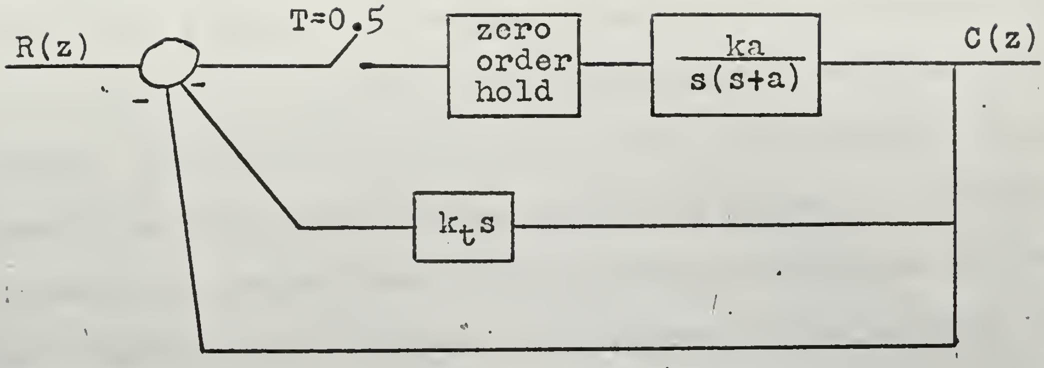


FIGURE 20a

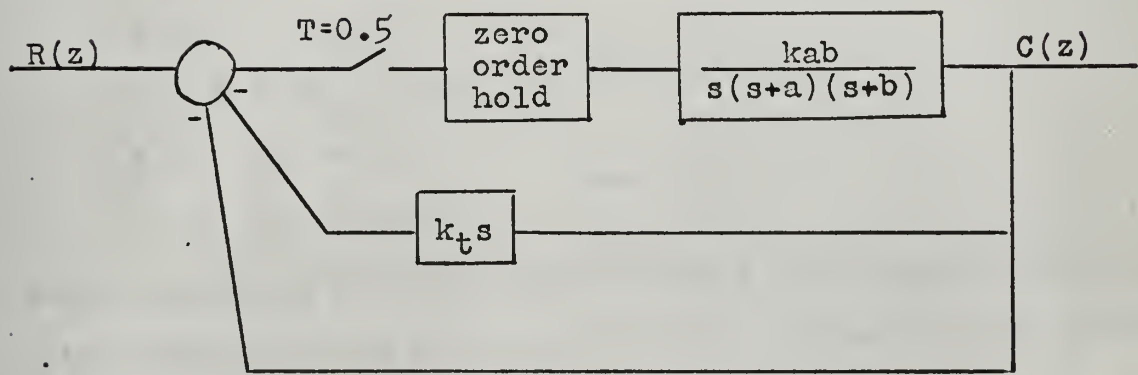


FIGURE 20b

on the distance that pole b is "removed" but additionally it depends on gain values k and k_t. A knowledge of when second order behavior prevails as a function of k, k_t, and b is of consequence in analyzing the overall system. Furthermore, if k and k_t are used to adjust overall system performance - for what range of values can these adjustments be treated as adjustments to a second order system and when must the third pole be considered? These questions are conveniently answered on the parameter plane, shown in figure 21, where it is assumed that nominally these poles are adjusted to a zeta value of 0.5. (If additional ζ values are of interest these contours are also mapped into the parameter plane). Proximity of second and third order $\zeta = 0.5$ curves denotes approximately equal characteristic equation roots for parameter settings in these regions; consequently the third order system is closely approximated by its second order counterpart. Figure 21 shows such approximations to be valid for

$$\left.\begin{array}{l} 0.6 < \alpha \triangleq k < 1.6 \\ 0 < \beta \triangleq kk_t < 0.6 \end{array}\right\} \quad \text{for } b = 10$$

$$\left.\begin{array}{l} 0.6 < \alpha < 1.2 \\ 0 < \beta < 0.35 \end{array}\right\} \quad \text{for } b = 5$$

Notice that even a 10:1 ratio between poles a and b does not justify second order approximation for $\alpha > 2.4$ or $\beta > 1.5$. In figure 22 the effect of these approximations on bandwidth is shown.

The author cautions against haphazard extension of these approximations. For instance, proximity of ζ - curves for a second and fourth order system would not be sufficient to guarantee the validity of a second order approximation. Not only must the location of the other two roots to the fourth order system be investigated but the region of proximity for the ζ - curves must be checked for a corresponding proximity of w_n values as well.

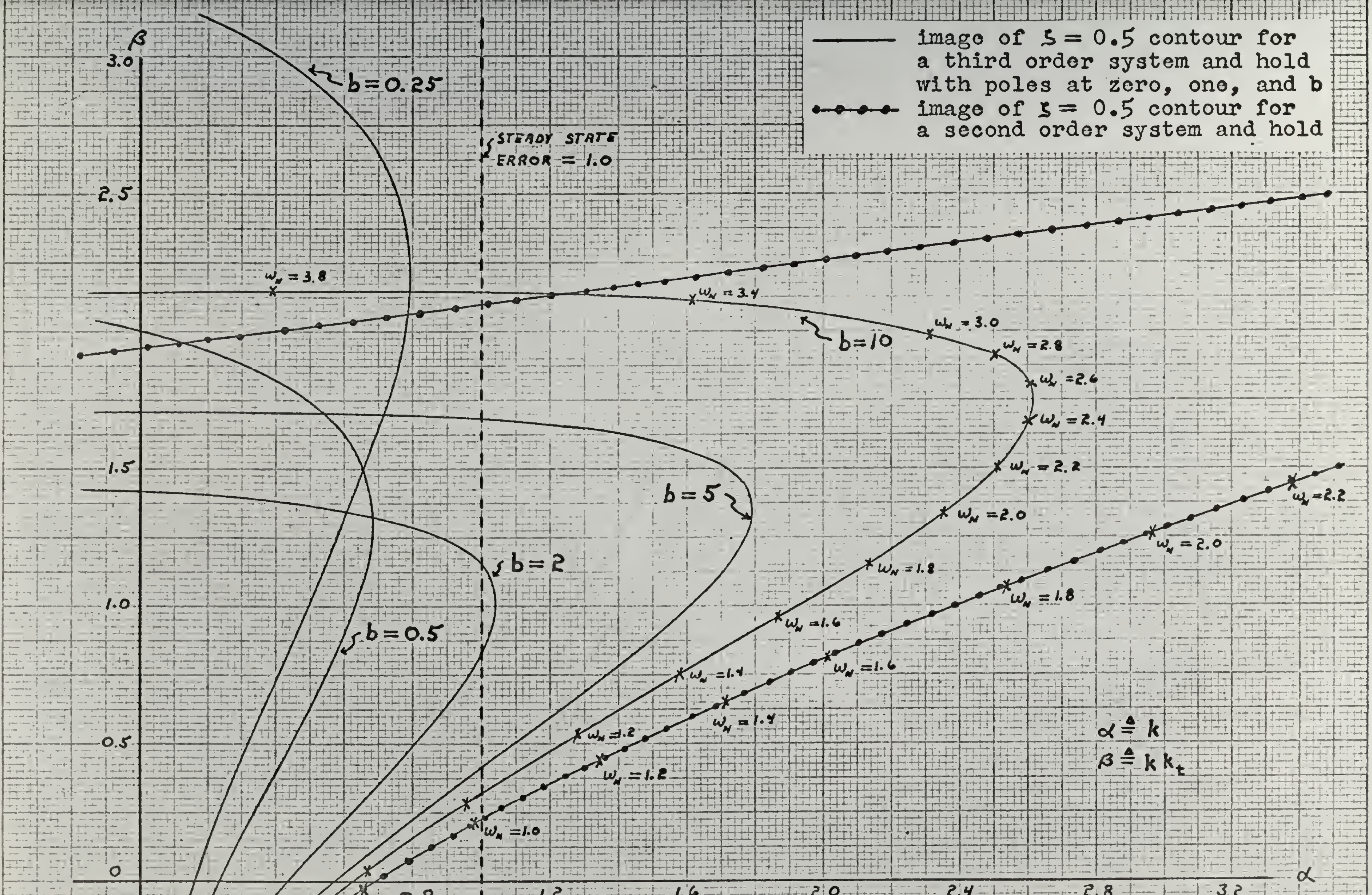


FIGURE 21

EFFECT OF A THIRD POLE ANALYZED ON THE PARAMETER PLANE

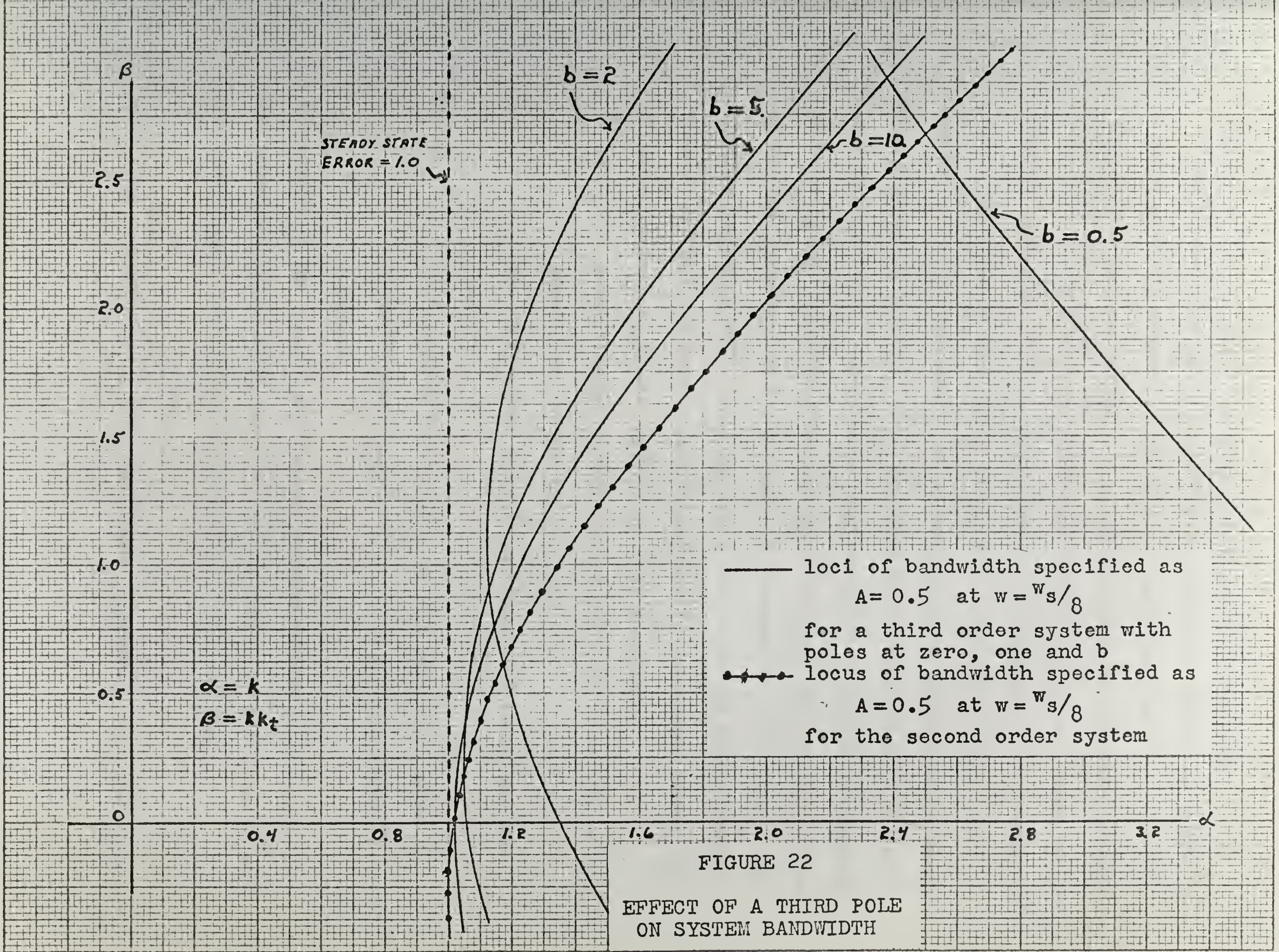


FIGURE 22

EFFECT OF A THIRD POLE
ON SYSTEM BANDWIDTH

Advanced Example of Compensation by Parameter Plane Technique

As mentioned in the introduction most graphical methods are seriously limited when called upon to display system performance measures as a function of more than one variable. In this respect parameter plane methods have definite advantages for performance indicators such as damping, settling time and bandwidth are inherently displayed as functions of two system variables. Furthermore, when only a single zeta value (or a single settling time or a single bandwidth) is of interest parameter plane methods competently display the effect of three variables on system performance. Even more variables can be considered without undue effort as will be illustrated in designing compensation for the system of figure 23. Here the designer has four parameters at his disposal, the system gain, the amount of derivative feedback, and the pole and zero locations of the single section filter. Settings for these parameters which allow the system to meet the following performance specifications are to be determined.

1. Maximum overshoot must be less than 1.25 or zeta is to be equal to or greater than 0.4.
2. The steady state error to unit ramp inputs must be less than 0.25.
3. The bandwidth (-3 db point) of this system must occur at a radian frequency greater than π to allow sufficiently fast response.

Examination of the closed loop transfer function reveals its suitability for parameter plane analysis with α and β capable of definition in any of several ways. The curves illustrated result from the choice

$$\alpha = k, \qquad \beta = kp/q$$

with p and k_t considered as third and fourth parameters. Initially k_t is

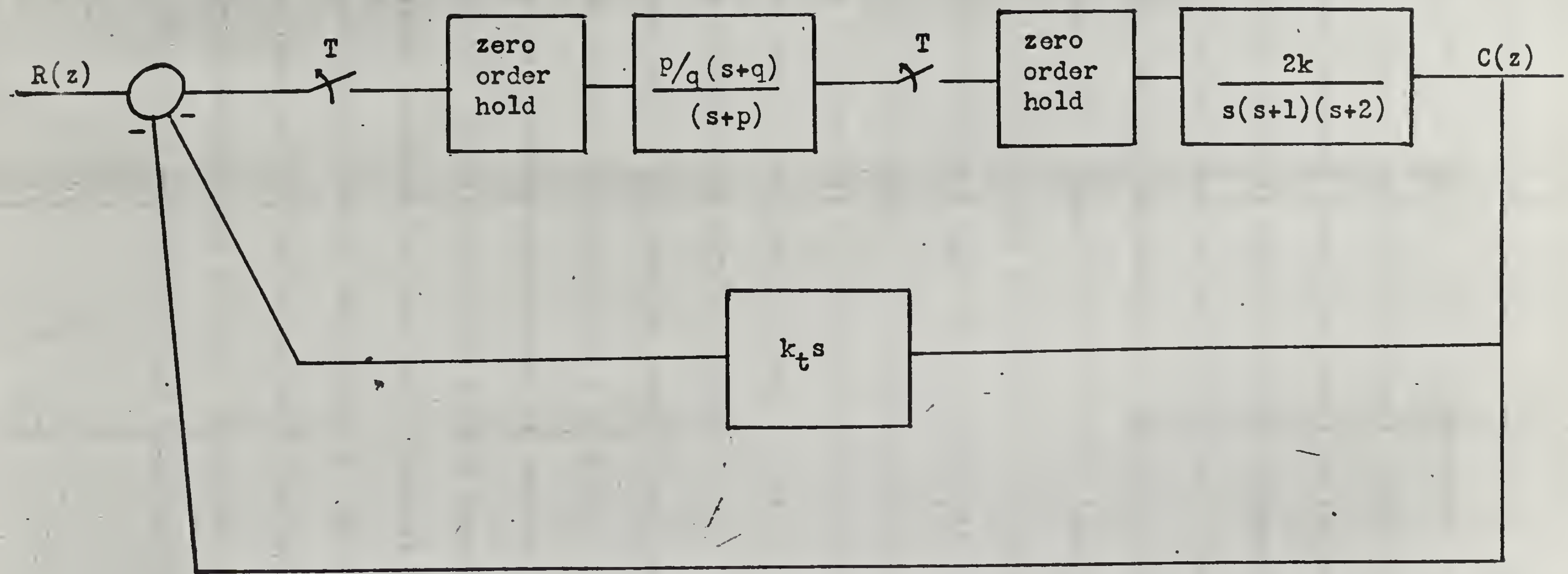


FIGURE 23

equated to zero and the efficacy of lag and lead compensation is investigated. The effect of such compensation on stability and bandwidth, as displayed in a parameter plane format, is shown in figures 24 and 25 respectively. That neither lag nor lead compensation is capable of rendering a satisfactory system is immediately evident. The intersection of the parameter plane region denoting stability with that region providing a satisfactory steady state error is empty for lead compensation whereas the intersection of those regions denoting stability with those denoting acceptable bandwidth is empty for lag compensation. However the parameter plane display does indicate that a "good location" for a lead section pole is around $s = -8$ and that a lag pole might profitably be located at $s = -0.05$.

The effect of derivative feedback in combination with lead and lag filters can now be investigated. The damping curves of figure 26 and the bandwidth curves of figure 27 display the results for various values of derivative feedback used in conjunction with a lead filter (or more accurately a single section filter whose pole is located at $s = -8$). Figure 28 displays an analogous result for feedback used in conjunction with a single section filter whose pole is located at $s = -0.05$.

In both instances system performance is enhanced by employment of some derivative feedback. In this case lag filters provide the easiest way to improve steady state accuracy but even with feedback their bandwidth limitations (slow rise times) cannot be overcome. On the other hand lead filters readily provide acceptable bandwidth and with derivative feedback both damping and error specifications can be met. Intersection of acceptable damping, error, and bandwidth regions indicate parameter values of

$$k = \alpha = 5.0 \qquad k_t = 0.5$$

$$p = 8.0 \qquad q = \frac{\alpha}{\beta} p = 3.0$$

should suffice.

[illegible]

[illegible]

[illegible]

[illegible]

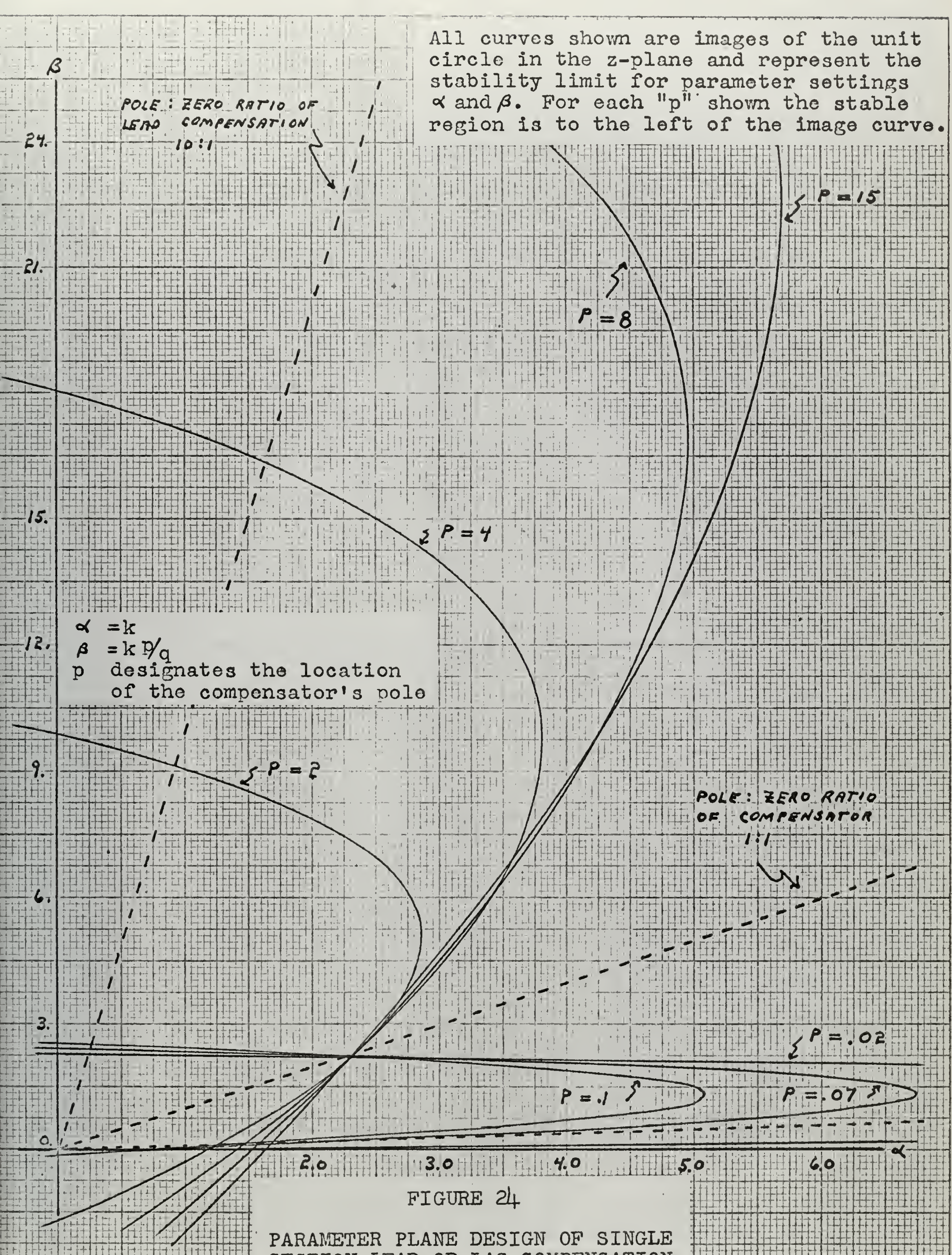


FIGURE 24

PARAMETER PLANE DESIGN OF SINGLE SECTION LEAD OR LAG COMPENSATION

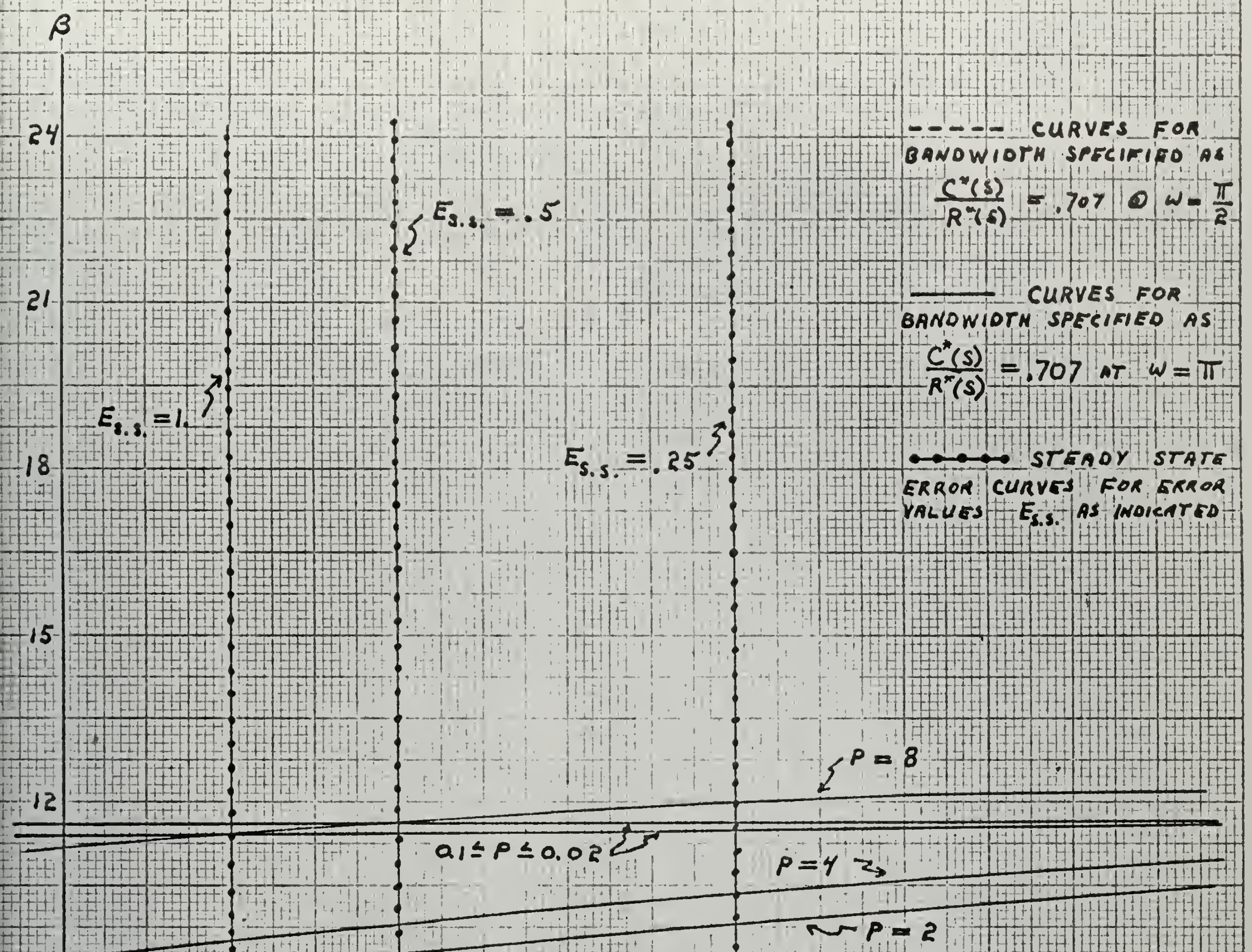


FIGURE 25

THE EFFECT OF THE POLE LOCATION OF A SINGLE SECTION COMPENSATOR ON SYSTEM BANDWIDTH AND STEADY STATE ERROR

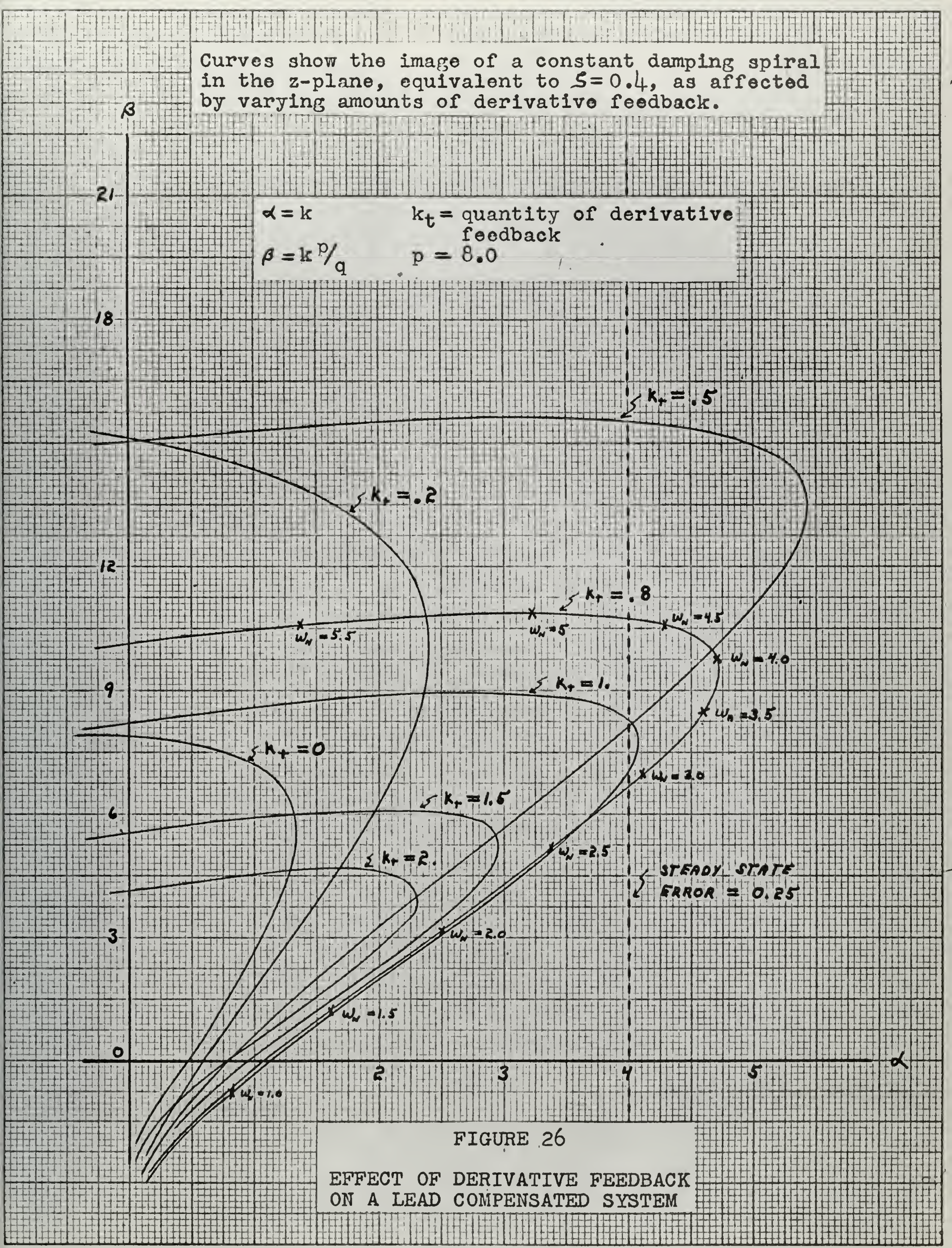


FIGURE 26

EFFECT OF DERIVATIVE FEEDBACK ON A LEAD COMPENSATED SYSTEM

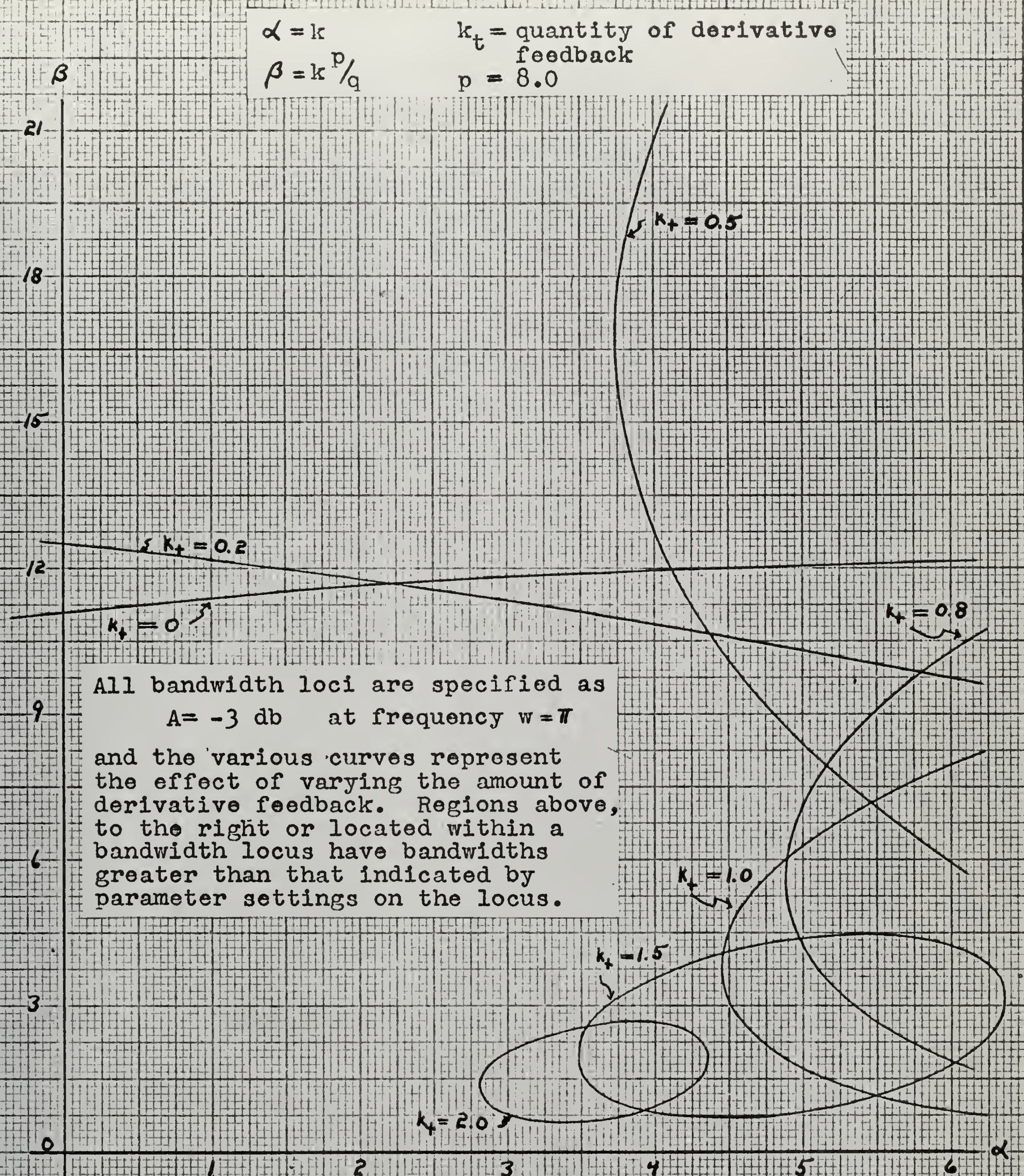


FIGURE 27

EFFECT OF DERIVATIVE FEEDBACK ON THE BANDWIDTH OF A LEAD COMPENSATED SYSTEM

Curves show the image of a constant damping spiral in the z-plane, equivalent to $\zeta = 0.4$, as affected by varying amounts of derivative feedback.

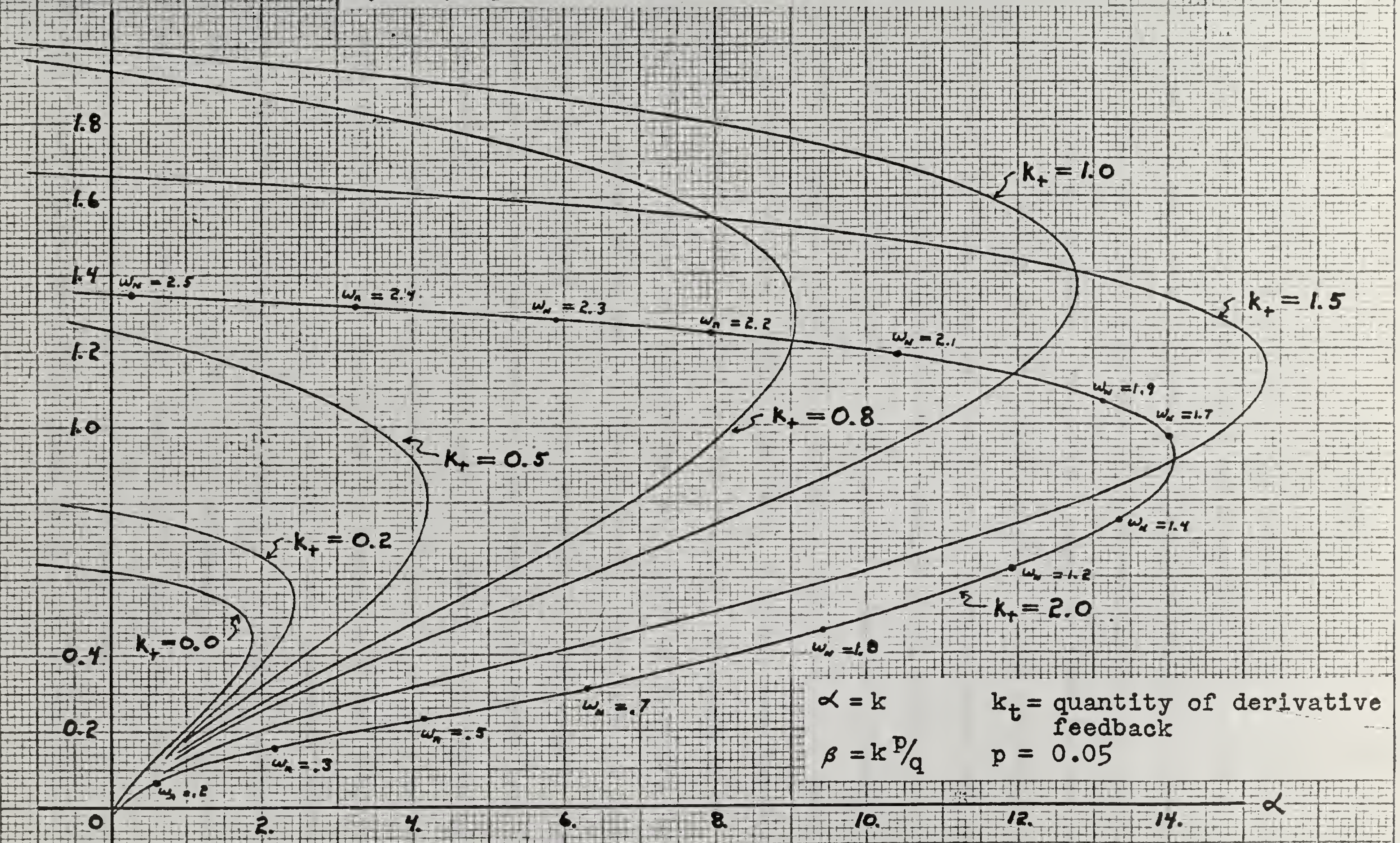


FIGURE 28

EFFECT OF DERIVATIVE FEEDBACK ON A LAG COMPENSATED SYSTEM

CHAPTER VII

The general availability of the digital computer has fostered revolutionary advances in servo-system and especially, in sampled data system design. However many of these new approaches completely discard rather than refurbish existing methods and thus void much of the experience gained by practising engineers. In contrast parameter plane methods retain the principle of design by such familiar specifications as damping and settling time but where classical methods could display the effect on only one variable on these specifications, parameter plane methods can display the simultaneous effect of two. This paper has extended parameter plane methods in two directions. First the requirement that parameters appear linearly in the characteristic equation was eased to permit consideration of those cases in which their product is also present. Extension in the second direction increased the information which could be displayed, and hence designed for, on the parameter plane to include bandwidth and steady state error specifications. This latter extension is particularly significant because it relates performance to the zeros of the closed loop transfer function which were previously ignored. Another novel capability of the parameter plane which was introduced in Chapter 5 is the simultaneous design of a dominant mode system and the guarantee of such dominance.

A great deal of further investigation into parameter plane methods in both theory and application still remains. To amplify questions typical of those to be resolved consider the consequences of allowing product terms to appear in the coefficients. Now points in the s- or z-plane no longer have unique images in the parameter plane and some points in the transform planes could conceivably have no real image at all. It is easy enough to postulate

that available parameters cannot place characteristic equation roots at these latter locations and that closed contours in the transform planes will result in consistent contours in the parameter plane. That is - these parameter plane contours are expected to be consistent with the interpretation expounded in Chapter 3. While such postulation seems reasonable enough for the product case, it certainly warrants a more rigorous framework if these methods are to be extended still further.

To the author's knowledge the bandwidth and error curves represent the first display of information on the parameter plane which is not directly obtainable from the characteristic equation. Research to discover other performance indicators capable of being displayed on the parameter plane should certainly be fruitful. In fact the author has obtained the equation governing loci of constant root sensitivity, but each point thereon must be computed by an iterative procedure and it is too unwieldy to be generally useful. Utilization of this auxiliary information in conjunction with that normally available on the parameter plane is illustrated in Chapter 6. Another facet of parameter plane methods illustrated by these problems is the relative ease with which the effect of a third variable parameter can be analyzed. This ease represents a significant improvement over the almost unworkable situation a three variable problem presents to root locus methods.

The author feels that parameter plane methods as enhanced by this paper should equal or exceed, for two parameter problems, those capabilities which the root locus brings to problems of one parameter. Furthermore, the reliance on familiar performance indicators minimizes prerequisite studies and enhances the experience of most engineers. These traits auger its more widespread appearance.

BIBLIOGRAPHY

1. Demetry, James. "Linear Control System Optimization Using a Model-Based Index of Performance", Unpublished Ph.D. dissertation, United States Naval Postgraduate School, 1964.

2. Kuo, B. C. Analysis and Synthesis of Sampled Data Control Systems, New Jersey, 1963.

3. Mitrovic, D. "Graphical Analysis and Synthesis of Feedback Control Systems, I Theory and Analysis, II Synthesis", AIEE Transactions Part II Applications and Industry, Vol. 77, 1958.

4. Niemark, J. I. "About Determination of Parameter Values which Insure Stability of Automatic Control Systems", Automatika i Telemakhanika, Number 3, 1948.

5. Numakura, T. and Miura, T. "A New Stability Criterion of Linear Servomechanisms by a Graphical Method", AIEE Transactions, Part II Applications and Industry, Vol. 76, 1957.

6. Ohta, T. "Mitrovic's Method - Some Fundamental Techniques", Unpublished research paper no. 39, United States Naval Postgraduate School, 1964.

7. Siljak, D. "Feedback Control Systems in the Parameter Plane, Parts I, II, & III", IEEE Transactions on Applications and Industry, Vol. 83 no. 75, 1964.

8. Thaler, G. J. and Brown, R. G. Analysis and Design of Feedback Control Systems, New York, 1960.

9. Lindorff, D. P. "Application of Pole-Zero Concepts to Design of Sampled-Data Systems", IRE Transactions on Automatic Control, Dec. 1959.

APPENDIX

Some Self-Explanatory Parameter Plane Curves

These self-explanatory curves should serve to relate the familiar second order system to some parameter plane representations of it. A knowledge of the general pattern of such relations will serve the reader when interpreting parameter plane curves for more complex systems.

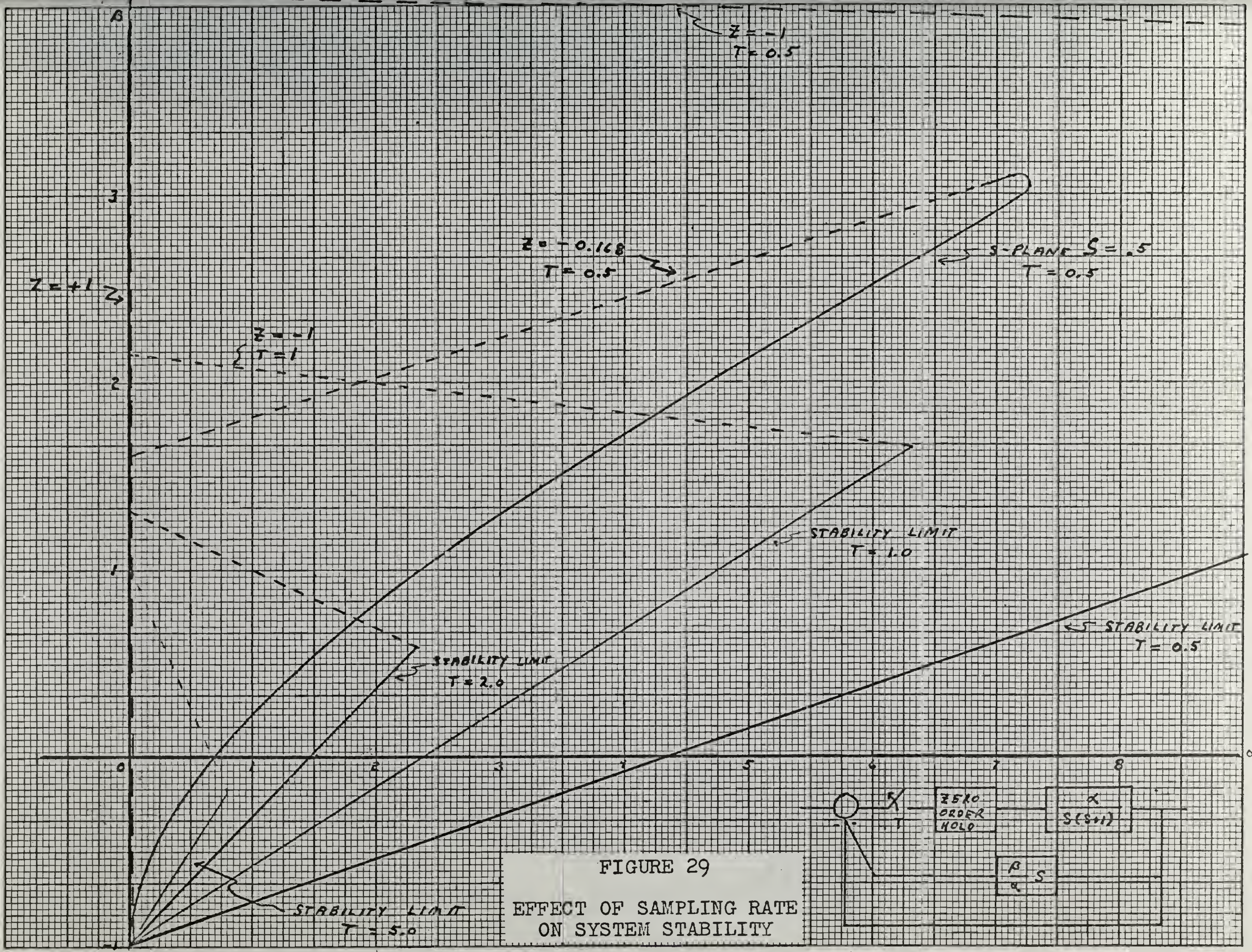


FIGURE 29

EFFECT OF SAMPLING RATE ON SYSTEM STABILITY

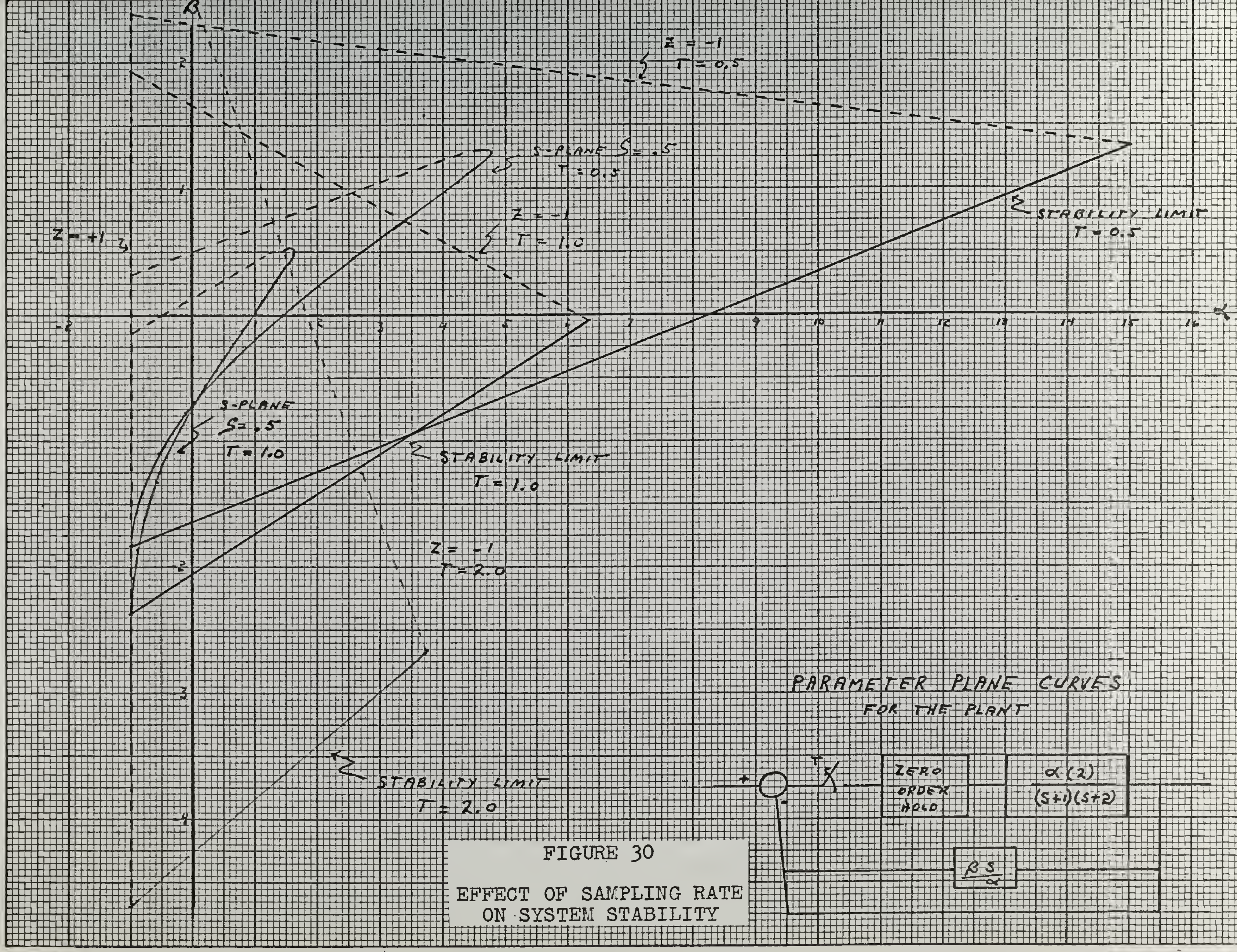


FIGURE 30

EFFECT OF SAMPLING RATE
ON SYSTEM STABILITY

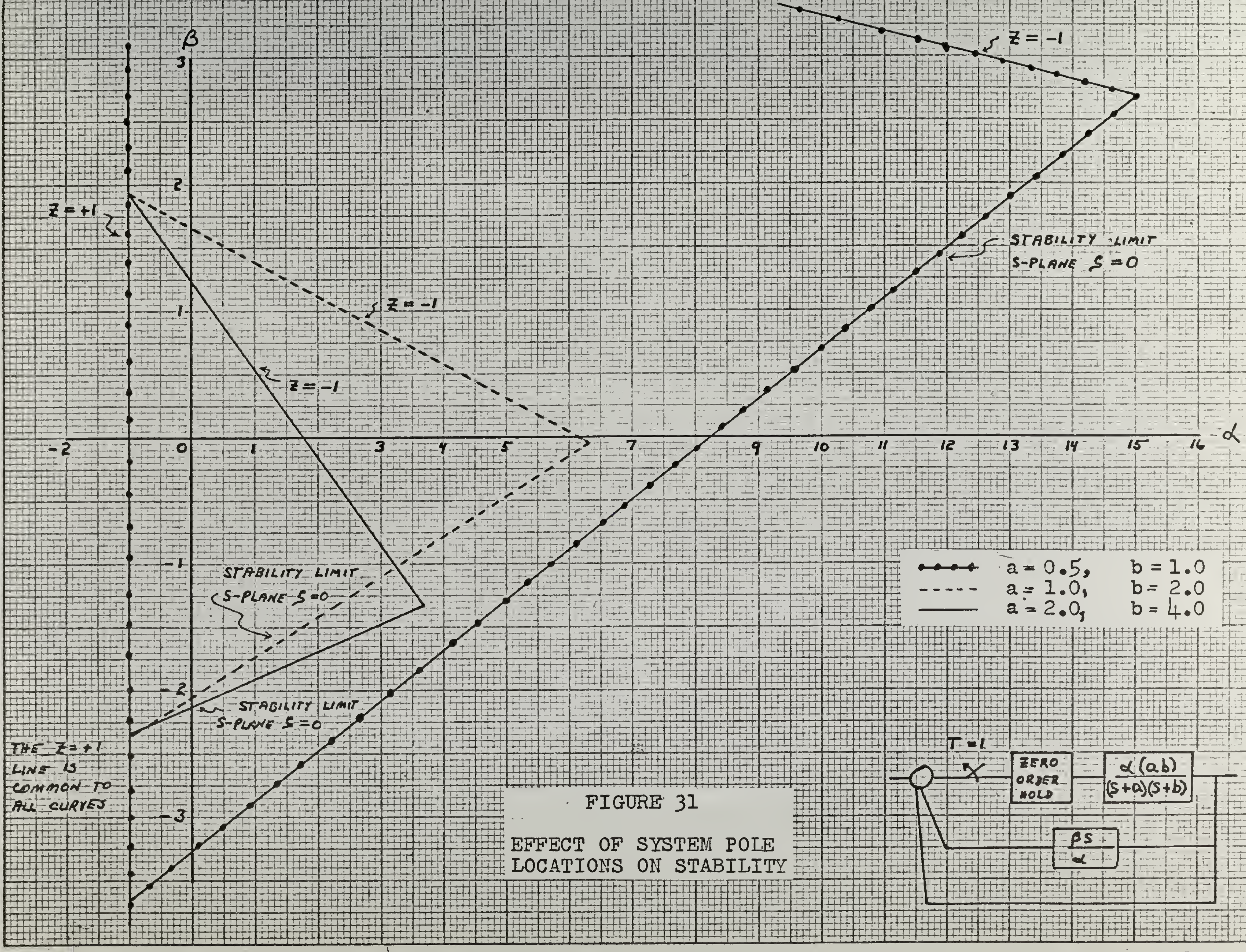


FIGURE 31

EFFECT OF SYSTEM POLE LOCATIONS ON STABILITY